# Mountain Goat Genetic Structure, Molecular Diversity, and Gene Flow in the Cascade Range, Washington

by

Andrew J. Shirk

Accepted in Partial Completion
of the Requirements for the Degree of
Master of Science

_____

Moheb A. Ghali, Dean of the Graduate School

ADVISORY COMMITTEE

_____

Chair, Dr. David O. Wallin

_____

Dr. Clifford G. Rice

_____

Dr. Kenneth I. Warheit

_____

Dr. Michael J. Medler

**MASTER'S THESIS**

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the nonexclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Signature_____

Date_____

# Mountain Goat Genetic Structure, Molecular Diversity, and Gene Flow in the Cascade Range, Washington

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

by
Andrew J. Shirk
May 2009

# ABSTRACT

Anthropogenic landscape changes have fragmented populations within virtually all taxonomic groups, resulting in impacts across a wide range of biological processes. Among these, increased genetic isolation reduces gene flow among subpopulations and genetic diversity within subpopulations. These effects have been positively correlated with extinction risk, making an understanding of population structure and genetic diversity critical to the conservation of small populations in fragmented landscapes. The classical population genetics approach to quantify genetic diversity has been to assign sampled individuals to discrete subpopulations, and then calculate indices of genetic diversity assuming panmixia within otherwise isolated subpopulations. These assumptions do not account for other mechanisms of genetic isolation such as isolation by distance (IBD) or isolation by landscape resistance (IBR), which result in clinal population structures, and can therefore lead to misinformed conclusions regarding gene flow and genetic diversity. In chapter 1, I explore the potential for both discrete and clinal models to capture the genetic structure and diversity of mountain goats (*Oreamnos americanus*) in the Cascade Range, Washington. This population has been greatly depleted, apparently by harvest, and now inhabits a landscape highly altered by anthropogenic factors, including potential barriers that may subdivide the population into discretely bounded panmictic subpopulations. The geographic scale and complexity of this landscape, however, may favor IBD or IBR. With discrete models, I found that an interstate highway corridor fragmented the population into two subpopulations, each exhibiting low genetic diversity. However, with clinal models, including a novel approach to

identify gradients of genetic diversity, I found that the assumption of panmixia within each subpopulation was not supported, and that discrete models masked substantial heterogeneity in genetic diversity. I discuss the implications of these findings for conservation of species in fragmented landscapes.

An understanding of genetic isolation within the study area will improve our ability to identify landscape factors that limit gene flow, resulting in informed management decisions and more effective strategies to improve gene flow. In Chapter 2, I explore the relative importance of barriers, distance, and landscape resistance to genetic isolation among mountain goats in the Cascades. While isolation due to distance or barriers is conceptually easy to model, models of isolation by landscape resistance (IBR) based on expert opinion are highly subjective. Here, I employed a causal modeling framework to rigorously test the assumptions of an expert opinion landscape resistance model with multiple alternative IBR hypotheses. I used a novel, multivariate approach to model selection as well as circuit theory to quantify landscape resistance. I found that an alternative IBR hypothesis was more highly correlated ($r = 0.723$) with genetic isolation in the Cascades than the expert opinion model, or other models based on isolation by barriers or distance. This IBR model indicated that an interstate bisecting the range, highways, development, low elevation valleys, and water bodies limit connectivity, and suggests modern landscape changes have greatly reduced gene flow. I conclude by discussing the implications of this modern landscape for the recovery and persistence of mountain goats in the Cascade Range, Washington.

# ACKNOWLEDGEMENTS

This work was part of a collaboration involving the Washington Department of Fish and Wildlife (WDFW), Western Washington University (WWU), the National Park Service (NPS), and the U.S. Forest Service (USFS). Funding for this project was provided by WDFW, the NPS, and grants from the WWU Office of Research and Sponsored Programs, Huxley College of the Environment, Seattle City Light, the Mountaineers Foundation, the Mazamas, and the Aquatic Lands Enhancement Account. I would like to thank my advisor, David Wallin, and my committee members Clifford Rice, Kenneth Warheit, and Michael Medler for their mentorship, expertise, and support. Additional thanks is due Clifford Rice for the training he provided on several forays into the field. I am also grateful to Kenneth Warheit and Cheryl Dean for performing all laboratory work for this project, including DNA isolation and genotyping. Thanks also to Samuel Cushman, USFS, for his expertise and guidance. Finally, I thank the interns, Brian Buma, Eric Fox, Patrick Haggerty, Cody Hoel, Tyler Klein, Brett Matulis, Patrick Migas, Megan Scherer, Andrea Swanson, and John Swensen, who assisted me in collecting genetic samples.

TABLE OF CONTENTS

ABSTRACT ..... iv

ACKNOWLEDGEMENTS ..... vi

LIST OF TABLES ..... ix

LIST OF FIGURES ..... x

PREFACE ..... 1

CHAPTER 1 ..... 1

INTRODUCTION ..... 4

METHODS ..... 7
Study area ..... 7
Sample collection ..... 8
Genotyping ..... 9
Population structure ..... 10
Genetic diversity and F-statistics ..... 12
Migration ..... 13

RESULTS ..... 14
Sample statistics ..... 14
Population genetic structure ..... 14
Population differentiation and migration ..... 16
Genetic Diversity ..... 17

DISCUSSION ..... 18
Discrete vs. clinal models of population structure ..... 18
Discrete vs. clinal models of genetic diversity ..... 19
Implications of model choice for conservation ..... 21

CHAPTER 2 ..... 30

INTRODUCTION ..... 30

METHODS ..... 34
Study area ..... 34
Sample collection, and genotyping ..... 34
GIS data ..... 35
Quantifying genetic distance ..... 35
Modeling IBB and IBD ..... 36

Modeling IBR ..... 36
Evaluating the relative support for IBB, IBD, and IBR ..... 37
Univariate IBR model optimization ..... 37
Multivariate IBR model optimization ..... 41
Predicting gene flow ..... 42

RESULTS ..... 42
Univariate optimization ..... 42
Multivariate optimization ..... 43
Relative support for IBB, IBD, IBR ..... 44
North-south landscape resistance and gene flow ..... 45

DISCUSSION ..... 46

LITERATURE CITED ..... 65

APPENDIX ..... 73
Formulae ..... 73
Microsatellite markers ..... 74

# LIST OF TABLES

## Chapter 1

Table 1. Genetic diversity indices for mountain goat subpopulations.....23

Table 2. Pairwise Θ values among mountain goat subpopulations.....24

## Chapter 2

Table 1. Resistance values for landscape factors influencing gene flow.....53

Table 2. Mantel and partial Mantel correlations for models of genetic isolation.....54

# LIST OF FIGURES

## Chapter 1

Figure 1. Map of the study area.....25

Figure 2. STRUCTURE and Geneland analysis of population genetic structure.....26

Figure 3. Number of subpopulations supported by STRUCTURE and Geneland analysis....27

Figure 4. Spatial principle components analysis of population genetic structure.....28

Figure 5. Moving window analysis of genetic diversity.....29

## Chapter 2

Figure 1. Map of the study area.....55

Figure 2. Univariate optimization of landscape resistance due to escape terrain.....56

Figure 3. Univariate optimization of landscape resistance due to roads.....57

Figure 4. Univariate optimization of landscape resistance due to landcover.....58

Figure 5. Univariate optimization of landscape resistance due to elevation.....59

Figure 6. Multivariate optimization of landscape resistance due to roads.....60

Figure 7. Multivariate optimization of landscape resistance due to elevation.....61

Figure 8. Multivariate optimization of landscape resistance due to landcover.....62

Figure 9. Multivariate optimization of landscape resistance due to roads (round 2).....63

Figure 10. Map depicting the most supported model of genetic isolation and gene flow.....64

# PREFACE

One quarter of all mammalian species are now listed in the three highest categories of extinction risk (critically endangered, endangered, and threatened) as defined by the International Union for Conservation of Nature (IUCN 2008). The IUCN attributes this extinction risk to three primary factors. Unsustainable levels of harvest have resulted in the depletion of many mammalian species, and the surviving remnant populations are often small and extinction prone. In addition, habitat destruction, commonly in the form of deforestation, conversion of habitat for agriculture, or urbanization, has greatly reduced the range of nearly all threatened mammalian species. Finally, partially as a result of extensive habitat loss, many of these species exist in populations that are highly fragmented, a condition which, as this thesis explores, also contributes to extinction risk.

Mountain goats (*Oreamnos americanus*) are currently listed as a species of "least concern" (IUCN 2008). Their low risk of global extinction is due primarily to their limited exposure to harvest, habitat loss, and habitat fragmentation over much of their native range in North America (Festa-Bianchet and Cote 2007). The limited impact of these factors is largely a function of the extreme isolation between most mountain goat populations and human habitats – they live in some of the most remote mountain ranges on Earth, and within those ranges, they are habitat specialists adapted to living in the steepest and most inaccessible environments. The low global risk of extinction among mountain goats, however, is not to say that locally, some populations in North America are without risk of extinction. Indeed, the population in the Cascade Range of Washington, which is the subject of this thesis, has been greatly depleted by harvest while, at the same time, modern landscape changes may

have resulted in habitat fragmentation. As a result, the viability of this population is now a management concern.

While the IUCN is focused primarily on global extinctions, the impact of local extinctions (extirpation) within any particular ecosystem is manifest regardless of whether the species is also extinct elsewhere. For example, the fact that grey wolves (*Canis Lupus*) are abundant in Canada and Alaska does not change the local effects arising from their extirpation over most of their former range throughout the contiguous United States. Extirpations and, more generally, loss of biodiversity can have profound effects on ecosystem structure and function (Naeem et al. 1994, Letters 2002, Hooper et al. 2005, Cardinale et al. 2006). The loss of large mammals in particular may have a disproportionate effect on ecosystem processes, as such species often play keystone roles regulating the abundance and distribution of other species, with additional effects arising through trophic cascades (Paine 1969, Mills et al. 1993). For example, reintroduction of wolves into the Yellowstone ecosystem altered the behavior and abundance of the ungulate species there, which, in turn, resulted in a restructuring of the plant community (Ripple and Beschta 2003). While the loss of mountain goats from alpine communities has not been studied, their role in supporting populations of predators such as grey wolves (Fox and Streveler 1986) as well as their effects on plant communities (Pfitsch 1981) would suggest this species plays an important role in the alpine ecosystems in which they have co-evolved for millennia.

The goal of this study was to investigate the potential for past harvest and modern habitat fragmentation in the Cascade Range, Washington to result in loss of genetic diversity, the effects of which may impair the recovery and long-term viability of the local mountain goat population. In Chapter 1, I delineated this population's genetic structure in order to

identify anthropogenic factors that might fragment the population. Within that structure, I quantified current levels of genetic diversity as well as the potential of the population to retain genetic diversity in the future. In Chapter 2, I modeled genetic isolation and gene flow within the Cascades to evaluate connectivity given the modern landscape. Together, the results and discussion presented in these chapters offer insight into the ability of this population of mountain goats to persist in the future, and suggest means by which wildlife managers may improve the population's long-term viability.

# CHAPTER 1

## The Genetic Structure and Molecular Diversity of Mountain Goats in the Cascade Range, Washington

### INTRODUCTION

Harvest, habitat loss, and habitat fragmentation are important factors shaping the current abundance and distribution of mammals (Scribner 1993, Ceballos and Ehrlich 2002, Wiegand et al. 2005). Small remnant populations inhabiting fragmented landscapes are susceptible to extirpation by environmental stochasticity and Allee effects (Lande 1998a, Dennis 2002), as well as several genetic processes (Lacy 1997, Gaggiotti 2003, Keyghobadi 2007). Among the genetic factors influencing population persistence, inbreeding depression (Keller and Waller 2002, Roff 2002, Reed et al. 2007) has been shown to strongly affect extinction risk in wild populations (Crnokrak and Roff 1999, Hedrick and Kalinowski 2000, Tanaka 2000, Brook et al. 2002, O'Grady et al. 2006). Small populations are also susceptible to loss of allelic diversity through genetic drift, increasing the risk of reduced evolutionary potential and random fixation of deleterious alleles (Lacy 1997, Willi et al. 2006). Lost fitness due to these genetic processes may promote additional population decline, further eroding genetic diversity and fitness until ultimately, the population is extirpated in what has been termed an extinction vortex (Gilpin and Soulé 1986, Lacy 1997, Tanaka 2000, Fagan and Holmes 2006).

To ascertain whether a population is susceptible to extirpation by genetic factors, an understanding of genetic structure (to infer the extent of genetic isolation) and genetic diversity (in terms of inbreeding and heterozygosity) is required. In classical population

genetics, population structure is either arbitrarily defined using discrete boundaries or, more recently, molecular methods are used to assign individuals to discrete subpopulations. In either case, genetic diversity indices are then calculated for all individuals within each subpopulation. This approach assumes that strong dispersal barriers subdivide a population into otherwise panmictic subpopulations. While this may be appropriate for some species in certain landscapes, this assumption does not account for alternative mechanisms of genetic isolation, including isolation by distance (Wright 1943) or isolation by landscape resistance (Cushman et al. 2006, McRae 2006). These latter two mechanisms both give rise to clinal populations, where genetic structure and diversity are characterized by gradients rather than discrete boundaries. If isolation by barrier is not the primary driver of genetic isolation within a population, assuming discrete boundaries and panmictic subpopulations may lead to erroneous estimates of genetic diversity, with concomitant adverse effects on the ability of wildlife managers to assess the role genetic factors may play in population viability.

In this study, I sought to gain insight into this issue by comparing the genetic structure and diversity of a mountain goat (*Oreamnos americanus*) population in the Cascade Range, Washington, using both discrete and clinal methods. In 1961, the Washington Department of Game (now Washington Department of Fish and Wildlife; WDFW) estimated this population to number about 8500 (excluding large populations within Mount Rainier National Park and Yakama Indian Nation lands, where no surveys were conducted). More recent WDFW surveys estimate 2512 (95% CI: 2137-2868) remain, including Mount Rainier National Park and all tribal lands (Rice, WDFW, unpublished data). A recent study provides evidence that the predominant cause of this decline is likely the high level of harvest that began in the 1940's and peaked in the 1960's (Rice and Gay 2009). However, despite an

almost complete ban on harvest during the past two decades, recovery has not been apparent in most areas of the Cascades, and large areas of formerly occupied habitat remain sparsely populated. This raises the possibility that genetic factors may play a role in the recovery of this population.

In addition to the effects of past harvest, the modern landscape may also affect the viability of mountain goats in the Cascade Range. Numerous forest service roads, several major state highways, and one interstate highway now cross the range. Several ski resorts and residential areas have been developed near major passes, and timber harvest has altered montane forests surrounding alpine mountain goat habitat. This modern Cascades landscape may act in conjunction with historic landscape factors to reduce migration between habitat patches and thereby fragment the population into smaller subpopulations exhibiting reduced gene flow. In particular, major highways have been shown to be instrumental factors contributing to fragmentation among a number of vertebrate species (Forman 2003, Coffin 2007). In the Cascades, Interstate 90 (I-90), the most heavily used roadway in the state, spans the entire range, and may therefore act as a strong barrier to north-south gene flow, effectively dividing the Cascades into two discrete subpopulations. Alternatively, the geographic scale and complexity of the modern landscape may favor isolation by distance or isolation by landscape resistance, resulting in a clinal population structure.

The potential for both discrete and clinal structure among mountain goats in the Cascades makes this population an ideal opportunity to compare classical discrete population genetics methods of assessing genetic structure and diversity with clinal methods. In this study, I used these two alternative approaches, including a novel means of quantifying spatial gradients of genetic diversity, to provide a contrasting but potentially complimentary

perspective that may improve our understanding of genetic factors affecting the recovery and long-term viability of mountain goats in the Cascade Range, Washington.

## METHODS

### Study area

The study area includes approximately 36,500 $km^2$ of the Cascade Range of Washington, extending some 315 km from Mount Adams to the Canadian border (Figure 1). This region represents the southern extreme of the coastal North American population of mountain goats, and appears poorly connected to the nearest populations in Idaho and British Columbia. The landscape is generally mountainous and, at high elevations, includes rocky alpine ridges and summits, cliffs, alpine meadows, parkland, and glaciers. At lower elevations, heavily forested glacial valleys predominate. Elevation ranges from near sea level to almost 4400 m. Annual precipitation averages from as low as 70 cm on the eastern slopes to over 350 cm at some high elevation areas west of the crest. Numerous roads are found throughout the Cascades, the largest of which is Interstate 90, which runs east-west across the crest of the range through Snoqualmie Pass. In addition, three state highways run east-west across the range (two north and one south of I-90) and numerous other highways and secondary roads intrude towards the crest mainly from the lower elevation valleys to the east and west. A few developed areas exist within the study area. They include several small towns and agricultural areas situated in lower elevation valleys along the periphery as well as several ski resorts located near the major mountain passes.

## Historic translocations

A non-native mountain goat population was established on the Olympic peninsula in the 1920's from 12 individuals drawn from founder populations in Alaska and British Columbia. In an effort to eliminate this species from Olympic National Park (ONP), the National Park Service translocated several hundred mountain goats to areas throughout the western United States, including a total of 130 animals to 15 sites (Figure 1) within the Washington Cascades (Houston et al. 1991). Because these translocations have the potential to confound genetic analyses of the native Cascade mountain goat population, samples were collected from the ONP population to be used for identification of ONP-Cascades admixed genotypes.

## Sample collection

I obtained genetic samples from the Cascade and Olympic Ranges (Figure 1) by multiple means between 2003 and 2008. The National Park Service (NPS) contributed a total of 17 blood samples (8 from ONP, 6 from Mount Rainier National Park, and 3 from the North Cascades National Park) obtained from animals chemically immobilized for the purpose of affixing GPS collars. WDFW also contributed 41 blood samples from throughout the Cascade Range obtained from animals chemically immobilized to attach GPS collars. WDFW contributed an additional 68 tissue samples donated from hunters legally issued a permit to harvest mountain goats in the Cascades. I, in collaboration with WDFW, also collected 29 Cascades and four Olympic Range tissue samples by biopsy darting via helicopter or on foot. Lastly, I obtained one bone marrow and one muscle sample from carcasses found in the field (cause of mortality unknown). In total, I included 149 genetic samples of all types from the Cascades 12 tissue and blood samples from the Olympic Range

in this study. All procedures were approved by the Animal Care and Use Committee at Western Washington University and permitted by WDFW and the National Park Service.

## Genotyping

Genotyping was performed by the WDFW molecular genetics laboratory in Olympia, Washington. They extracted DNA from tissue, bone marrow, or blood with the commercial DNeasy® blood and tissue DNA isolation kit (Qiagen) according to the manufacturer's protocol. They eluted DNA from the column with 100 µl of elution buffer, and used PCR to amplify 19 previously characterized polymorphic microsatellite markers, using six separate uni- or multi-plex reactions (multiplex Oam-A [BM121, BM4107, BM6444, TGLA122], Oam-B [OarCP26, OarHH35, RT27], Oam-C [BM1818, BM4630, RT9, URB038], Oam-D [BM203, BMC1009, HUJ616, McM527], Oam-E [BM1225, BM4513, HEL10], and TGLA10; PCR conditions available from WDFW) determined to effectively amplify mountain goat DNA (Mainguy et al. 2005). PCR products were visualized using an ABI 3730 capillary sequencer (Applied Biosystems), and sized using the GeneScan 500-Liz size standard (Applied Biosystems) and Genemapper 3.7 (Applied Biosystems).

I tested for deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) with GENEPOP 3.4 (Raymond and Rousset 1995). For loci with fewer than four alleles, I used the complete enumeration method (Louis and Dempster 1987). In all other cases, I used the Markov chain method (Guo and Thompson 1992). I adjusted significance values for multiple comparisons with a Bonferroni correction (Rice 1989). I used the program MICROCHECKER 2.2.0 (Van Oosterhout et al. 2004) to screen for genotyping errors, including allelic dropout, null alleles, and stuttering that might hinder

detection of heterozygotes. I corrected null allele estimates for inbreeding (calculated for each STRUCTURE subpopulation) with Null Allele Estimator 1.3 (Van Oosterhout et al. 2006).

### Population structure

I used three methods (two discrete and one clinal) to assess population structure, each of which makes different assumptions about the data regarding HWE and LD, or the spatial location of the sample. Multiple methods were required because I sought to detect both discrete and clinal structure. Moreover, because the spatial location of any Cascade descendants of translocated ONP goats would be an artifact of their recent ancestor's translocation, I first needed to identify these individuals with a method that does not use the spatial location of the sample to influence assignments. For this I used STRUCTURE 2.2 (Pritchard et al. 2000) which performs Bayesian inference of the most likely number of populations sampled and assigns individuals to populations based on minimizing Hardy-Weinberg and linkage disequilibrium within populations. To achieve this, the program uses a Markov chain Monte Carlo (MCMC) simulation to estimate the posterior probability that the data fit the hypothesis of $K$ populations, $P(X|K)$. I tested values of $K$ ranging from 1 to 10 with 5 independent runs per test (admixture model with correlated allele frequencies, a 100,000 step burn-in followed by $10^6$ steps of data collection). I chose the best supported value of $K$ based on the second order rate of change ($\Delta K$) in probability described by (Evanno et al. 2005). The admixture model also calculates the fractional probability ($Q$) of individuals belonging to each population.

Including descendants of ONP goats translocated to the Cascades would potentially confound this analysis of population structure. While all Cascades samples had some background posterior probability of ONP membership, samples with low probability of ONP ancestry are unlikely to bias results, and their omission reduces sample size for analyses of the Cascades population. As a balance between bias and sample size, I required the posterior probability of ONP population membership to be greater than 0.25 in order to classify a Cascade sample as "ONP-Cascade admixed." I evaluated the native population structure of mountain goats in the Cascades without the confounding effect of translocations by excluding ONP and ONP-Cascades admixed samples and rerunning STRUCTURE on native Cascades samples with the same parameters and criteria as above.

In addition to STRUCTURE, I used Geneland 3.1.3 (Guillot et al. 2005) as a complimentary approach to assess population structure. Geneland also performs Bayesian inference with a MCMC simulation and a population model based on minimizing Hardy-Weinberg and linkage disequilibrium. However, Geneland allows inclusion of the spatial coordinates to inform the prior distribution. I excluded all ONP-Cascades admixed samples, and then assessed the support for 1 to 10 populations with a burn-in of 100,000 iterations followed by $10^6$ additional iterations, from which every $100^{th}$ observation was sampled to reduce autocorrelation of the parameter estimates. I used the uncorrelated allele frequency model, which has been shown to perform better at estimating $K$ than the correlated allele model (Guillot et al. 2005), as well as the option to filter null alleles. After inferring $K$, I then re-ran the Geneland simulation fixed at $K$ to derive the population membership assignments and the posterior probability maps.

As an alternative to population based methods, I also inferred population structure with spatial principle components analysis, sPCA (Jombart et al. 2008). This approach, implemented in the R package Adegenet (Jombart 2008), makes no assumptions regarding a population model, nor does it assign individuals to discrete subpopulations. Rather, sPCA ordinates genotypes along a continuum of relatedness taking into account both the variance between individuals and spatial autocorrelation. As such, sPCA is an ideal tool to spatially represent populations structured as a cline. I used a nearest neighborhood by distance connection network where samples were connected if between 0 and 60 km (few long distance dispersal events have been known to exceed this distance, (Stevens 1983, Festa-Bianchet and Cote 2007) I then tested the significance of the global and local scores with the Monte Carlo tests implemented in Adegenet.

### Genetic diversity and F-statistics

For each subpopulation identified by our discrete-method analysis of genetic structure (STRUCTURE and Geneland), I estimated observed heterozygosity - $H_o$, Nei's gene diversity - $H_e$ (Nei et al. 1978), inbreeding coefficient - $F_{IS}$, and Θ (Weir and Cockerham 1984), with FSTAT version 2.9.3.2 (Goudet et al. 2002). In addition, I tested the significance of the differences in $H_o$ and $H_e$ with the two tailed non-parametric test implemented in FSTAT with 10,000 permutations. I also calculated the effective number of alleles ($A_E$).

In addition, I calculated individual-based diversity indices within a moving window in a raster GIS environment. This novel approach does not assume discretely bounded subpopulations, and therefore has the potential to detect gradients of genetic diversity. I first generated a point layer containing all Cascades genetic samples (including ONP-admixed

samples). Each point was joined to a table of individual allele usage (each allele was scored as 0 for not used, 1 for heterozygous, and 2 for homozygous), which I used to calculate multilocus heterozygosity (% of heterozygous loci) - $H$, and $H_e$ in a circular moving window with a radius of 24 km. While I used a 60 km window in the sPCA analysis to model the maximum possible distance at which samples were connected, I chose a more typical dispersal distance of 24 km to reflect the diversity of individuals most likely to contribute to genetic diversity near the center of the window. I also calculated $A_E$, and $F_{IS}$ in the same manner. I specifically chose these diversity indices because they are corrected for sample size, however, I excluded pixels where fewer than five individuals were within dispersal distance in order to minimize bias in sparsely sampled regions.

### Migration

To assess the degree to which each subpopulation was isolated by barriers, I sought to estimate recent migration rates between all subpopulations. For this I used a Bayesian approach based on multilocus microsatellite genotypes as implemented in the software package BayesAss+ (Wilson and Rannala 2003). BayesAss+ makes fewer assumptions than estimators of long-term gene flow, including relaxing the requirement for populations to be in HW equilibrium. After a burn-in period of $10^6$ iterations, $2x10^6$ iterations were performed and sampled at a frequency of 1 in 2000. I set the delta value for allele frequency, migration rate, and inbreeding at 0.14, 0.07, and 0.13 respectively. These values were chosen such that the accepted number of proposed changes was approximately 50% of the total number of iterations, as recommended by the software authors.

## RESULTS

### Sample statistics

The 161 genotypes I analyzed were, on average, 97.4% complete. None of these were > 95% identical, indicating that each sample was a unique individual. I re-amplified and genotyped 4.5% of samples to estimate the error rate, which was 0%. The samples included in this study represent approximately 5% of the total population of Washington. Eighteen of the 19 loci genotyped were polymorphic. I excluded the monomorphic locus URB038 from all analyses. I found no evidence for LD among any of the loci after Bonferroni correction for multiple comparisons (alpha = 0.05), but I observed significant departures from HWE in 13 loci when all samples were evaluated as a single population. When a population is structured, however, disequilibrium is predicted by the Wahlund effect (Wahlund 1928). After dividing the data into populations that correspond with the STRUCTURE clusters, none of the loci showed significant departures from HWE after Bonferroni correction for multiple comparisons. No evidence for allelic dropout or stuttering was detected by MICROCHECKER; however, 7 loci (BM203, BMC1009, BM1818, HEL10, BM1225, TGLA10, HUJ616) showed significant homozygote excess, suggesting the presence of null alleles. After correcting for inbreeding the estimated frequency of null alleles was less than 0.03 for all 7 loci. I therefore elected to retain these loci for all analyses.

### Population genetic structure

The $\Delta K$ value decreased markedly after $K$=3 when all samples were included in the STRUCTURE assignment test, indicating the highest support for 3 populations (Figure 3a). These three clusters correspond geographically to the Olympic Range, the north Cascades,

and the south Cascades (Figure 2a). The boundary between the north and south Cascades coincides with the I-90 corridor. The ONP cluster contained all 12 individuals sampled from the Olympic range, each of which was assigned with near 100% probability (Figure 2b). Fourteen samples in the Cascades had a probability ≥ 0.25 (mean 0.45, maximum 0.67) of belonging to the ONP population and were therefore classified "ONP-Cascades admixed". Based on the probability of population membership (Figure 2b) it appears nearly all the ONP-Cascades admixture occurs in the south Cascades. Furthermore, aside from one sample (identified as a first generation migrant by BayesAss+), the south Cascades does not appear highly admixed with the north Cascades, however, the north Cascades contains numerous individuals with a high probability of south Cascades subpopulation membership. Rerunning STRUCTURE excluding the ONP and ONP-Cascades admixed samples yielded a $\Delta K$ peak at $K$=2, supporting the 2 Cascades population structure that was evident when all samples were included (Figure 3a).

The Geneland analysis, which excluded ONP and ONP-Cascades admixed samples, also strongly favored two Cascade populations ($K$=2 was the most frequent value along the simulation chain, Figure 3b) and similarly identifies the genetic discontinuity corresponding to the I-90 corridor (Figure 2a). A zone of lower probability of belonging to the north Cascades subpopulation exists near Glacier Peak and the east-central Cascades. This zone coincides with the location of several samples classified by STRUCTURE as having a higher probability of belonging to the south Cascades subpopulation than the north Cascades subpopulation.

I observed a significant global structure within the Cascades by sPCA analysis ($P$ < 0.00001) indicating positive spatial autocorrelation of genotypes at the scale of the study area

(Figure 4b, first eigenvalue). The test for local population structure (negative autocorrelation at the local scale) was also significant ($P = 0.034$), however, the magnitude of the first local eigenvalue was very small, indicating local structure is negligible. First global axis sPCA scores, interpolated to aid in visualization (Figure 4a), revealed a clinal population structure along a north-south axis, with the most negative values located in the far north, increasing towards zero approaching I-90, and then becoming positive south of I-90. In areas where sample density was low, the interpolated sPCA score may not accurately represent the local sPCA values.

### Population differentiation and migration

Global Θ (Weir and Cockerham, 1984; an estimate of $F_{ST}$) for all samples was 0.187 indicating significant population substructure. I calculated pairwise Θ values between all three STRUCTURE subpopulations, excluding the ONP-Cascades admixed samples (Table 2). Θ ranged from 0.112 to 0.417, and all pairwise comparisons were significant after Bonferroni correction for multiple tests (alpha = 0.05).

I estimated recent migration rates with BayesAss+ to be 0.028 per generation from the south Cascades to the north Cascades and 0.010 per generation in the opposite direction. Three samples in the north Cascades (collected from GPS collared goats in the Glacier peak/east-central Cascades area) that were identified by STRUCTURE as belonging to the south Cascades population were classified by BayesAss+ as offspring of first generation migrants originating from the south Cascades. One sample in the Goat Rocks area (south Cascades) was identified as a first generation migrant from the North Cascades. This sample was submitted by a hunter permitted to harvest a goat from Goat Rocks, however, a

migration to that area from the north Cascades would represent an extreme dispersal distance that also involved crossing I-90, raising the possibility the harvest location was incorrectly reported by the hunter.

### Genetic diversity

Based on discrete population calculations, diversity was variable among the two Cascades populations, the ONP population, and the ONP-Cascades admixed group. $H_O$ ranged from a low of 0.351 in the south Cascades (excluding ONP admixed goats) to a high of 0.489 within the ONP-Cascades admixed goats. $H_e$ varied from 0.394 in the south Cascades (excluding ONP-Cascades admixed goats) to 0.502 among ONP-Cascades admixed goats. $F_{IS}$ was lowest among ONP-Cascades admixed goats (0.026, $P = 0.2917$) and highest among ONP goats (0.187, $P = 0.0014$). $A_E$ ranged from a low of 1.92 in the ONP population to a high of 2.41 among ONP admixed individuals.

As an alternative to discrete population statistics, I calculated diversity indices ($H_O$, $H_e$, $A_E$, *and* $F_{IS}$) for the Cascades in a 24km dispersal radius moving window (Figure 5). The increased spatial resolution of the moving window analysis revealed high variability within the two Cascades subpopulations. $H_O$ ranged from 0.303 on the west side of Mount Baker to 0.503 near Glacier Peak. $H_e$ ranged from a low of 0.356 on the west side of Mount Baker to a high of 0.522 near Glacier Peak. $F_{IS}$ varied from a low of -0.084 in the Okanogan area to a high of 0.231 southeast of goat rocks. $A_E$ ranged from a low of 1.35 in the southeast Cascades to 2.42 near Glacier Peak.

# DISCUSSION

## Discrete vs. clinal models of population structure

In this study, I found that discrete population based methods and clinal population methods offered differing perspectives on the genetic structure and diversity of Cascade Range mountain goats. From a discrete population perspective, assignment tests using both STRUCTURE and Geneland revealed a discontinuity that corresponds closely to the location of the I-90 corridor. Interstate 90 is by far the largest east-west travel route in the state and receives high traffic volumes during the summer months when mountain goats are most likely to disperse. While no studies have been conducted regarding the behavior of mountain goats with respect to road traffic, this species generally avoids human habitats. As mountain goat deaths from traffic collisions have not been observed along I-90 (Washington State Department of Transportation, pers. comm.), the reduced migration across this boundary may be rooted in behavior rather than mortality. In addition to the interstate, three ski resorts, several housing developments, and a large reservoir have been constructed near where I-90 crosses the crest of the Cascades at Snoqualmie pass. This high elevation area contains the most escape terrain and alpine habitat of any stretch along I-90 and therefore may have historically been the best dispersal corridor linking the north and south Cascades. As such, development in this area may add to the barrier effect of the interstate and further reduce gene flow.

The high differentiation detected between the north and south Cascades subpopulations suggests I-90 strongly affects migration. From a discrete population perspective, therefore, the Cascade Range is divided into north and south subpopulations. Additional subdivision of the north and south Cascades subpopulations was not supported by

the STRUCTURE or Geneland analysis. When calculating genetic diversity indices based on this discrete population structure, one assumes genetic diversity within subpopulations is uniform (i.e. panmixia), and that no individuals from other subpopulations contribute to genetic diversity (i.e. isolation by barriers).

Multiple lines of evidence, however, suggest this discrete population perspective does not fully capture the genetic isolation present within the study area. The detection of a first generation migrant (possibly an artifact, as described in the results) and three offspring of first generation migrants by BayesAss+ indicates I-90 is not an absolute barrier to migration. This is supported by the degree of admixture between the two subpopulations detected in the STRUCTURE $Q$-plot. Additionally, the sPCA analysis revealed a north-south gradient of relatedness that suggests a clinal population structure, raising the possibility that isolation by distance or isolation by landscape resistance may act in conjunction with the barrier effect of I-90 to increase genetic isolation. In chapter 2, I use a landscape genetics approach to evaluate the relative support for these alternative hypotheses of genetic isolation.

### Discrete vs. clinal models of genetic diversity

Assuming isolation by barrier (due to I-90) and panmixia within each subpopulation, the estimated heterozygosity, inbreeding, and allelic diversity reveal genetic diversity within the Cascades is significantly lower compared to the large and functionally connected populations that inhabit the core mountain goat habitat of North America in Alberta and British Columbia (only data for the Caw Ridge, Alberta population has been published – see Table 1; however, a recently completed study by A. Shafer et al., University of Alberta, unpublished, reveals ten other core populations have similar $H_O$, $H_e$, $A_E$, and $F_{IS}$ values compared to Caw Ridge).

In the short term, the low heterozygosity I observed is particularly troubling, as this has recently been shown to be negatively correlated with survival among juvenile mountain goats (Mainguy et al. 2009). In the long-term, low allelic diversity over much of the Cascades relative to core populations in North America suggests a more limited evolutionary potential for this population.

The discrete subpopulation perspective of genetic diversity, however, masks considerable heterogeneity I detected with the moving window analysis. With this approach, I found that the central Cascades, particularly north of I-90, exhibited much higher levels of heterozygosity and allelic diversity as well as lower inbreeding than predicted by the discrete approach. Indeed, values in this area were comparable to the core North American populations. On the other hand, the northern and southern portions of the range exhibited lower heterozygosity and allelic diversity and higher inbreeding levels than predicted by the discrete approach. This finding is consistent with a gradient population structure, which would predict higher genetic diversity in the central cascades due to multiple avenues for gene flow from both the north and south. The low diversity levels observed in the far south Cascades is also consistent with the position of this area at the southern terminus of the coastal North American mountain goat population, where gene flow is limited to a single direction. Importantly, the lower heterozygosity and higher inbreeding levels identified by the moving window analysis suggest inbreeding depression may be a much more immediate threat to population persistence and recovery in these areas than would have been predicted by the discrete subpopulation-level analysis.

### Implications of model choice for conservation

Wildlife passages and translocation are two widely employed strategies to improve gene flow and therefore benefit population viability in fragmented landscapes. In the Cascades, for example, three wildlife passages potentially suited for mountain goats are planned for I-90 (noted on Figure 1). The spatial gradients of genetic structure and diversity I detected may have important implications to the success of these strategies that were not apparent in the discrete population analysis. Notably, the additional genetic differentiation detected by sPCA suggests that the north subpopulation in particular is not panmictic, raising the possibility distance or other landscape factors, in addition to I-90, may limit gene flow. If so, any increase in north-south gene flow due to wildlife passages may not affect diversity within the subpopulation uniformly. Rather, population genetic theory would predict gene flow through wildlife passages would raise allelic diversity and heterozygosity as well as lower inbreeding in areas proximal to these structures, but with diminishing impact along the same distance or landscape resistance gradient that shapes the population's clinal structure. The mechanism of genetic isolation may also influence the optimal placement of wildlife passages. Isolation by distance, for example, would suggest such structures should be located in regular intervals to minimize the distance to the nearest passage, while IBR would favor placement to minimize landscape resistance along north-south corridors. Under a barrier model of isolation, any genetic variation crossing I-90 would be predicted to spread uniformly, increasing diversity throughout the subpopulation regardless of where the passages are constructed.

In addition to wildlife structures, translocation between subpopulations may be used to artificially improve gene flow. Here too, the mechanism of isolation bears strongly on the

conservation strategy. To improve genetic diversity in an area isolated primarily by IBD or IBR, translocated animals would be most effectively placed at sites that minimize distance or landscape resistance, respectively, to the target area. By contrast, translocations anywhere within a subpopulation would be predicted to improve genetic diversity under a model of subpopulation panmixia. In the south Cascades, which appear panmictic from both a discrete and clinal population model perspective, I detected 14 descendants of ONP goats translocated in the 1980's. In the five or so generations since these translocations, the distribution of ONP-Cascade admixed animals still appears strongly tied to the release sites, however, with time, the unique ONP alleles may continue to spread throughout the south Cascades, supporting the assertion of panmixia. Importantly, the potential for translocation to improve genetic diversity among mountain goats is reflected in the high genetic diversity of these ONP-Cascade admixed animals (Table 1).

This study highlights the potential for clinal models such as sPCA and the moving window analysis of genetic diversity I introduce here to offer a complimentary approach to understanding genetic structure and diversity, leading to more informed assessments regarding the role genetic factors may play in population viability. Ultimately, the future of the Cascades population will be determined by the complex interaction of many factors, including population demographics, genetic structure, gene flow, and small population extinction dynamics. This study highlights the need for incorporating these factors, as well as management strategies such as wildlife passages and translocation, into a population viability assessment to determine which scenarios best support a sustainable population of mountain goats in the Cascade Range of Washington.

# TABLES

**Table 1.** Sample size ($n$), genetic diversity summary statistics[a], fixation indices ($F_{IS}$), and significance test for inbreeding for mountain goat populations in Washington and Alberta, Canada.

| | $n$ | $A_E$ | $H_o$ | $H_e$ | $F_{IS}$ | $F_{IS}$ $P$-value |
|---|---|---|---|---|---|---|
| North Cascades | 63 | 2.21 | 0.41 | 0.46 | 0.10 | 0.0007 |
| South Cascades + ONP [b] | 86 | 2.04 | 0.37 | 0.43 | 0.12 | 0.0007 |
| South Cascades – ONP [b] | 72 | 1.97 | 0.35 | 0.39 | 0.11 | 0.0007 |
| Olympic National Park | 12 | 1.92 | 0.37 | 0.46 | 0.19 | 0.0014 |
| ONP-Cascades admixed | 14 | 2.41 | 0.49 | 0.50 | 0.03 | 0.2917 |
| Caw Ridge, Alberta [c] | 215 | 2.34 | 0.55 | 0.53 | -0.03 | - |

[a] $A_E$ ,effective number of alleles; $H_o$ observed heterozygosity; $H_e$ gene diversity

[b]The south Cascades population is calculated twice, first by including the ONP-Cascades admixed samples and then with those samples excluded.

[c]Values for the Caw Ridge population are based on 16 of the 18 loci used to calculate diversity for the Washington populations. $F_{IS}$ and $A_E$ were calculated from the originally reported $H_o$ and $H_e$ (Mainguy 2005). The $F_{IS}$ $P$-value for this population has not been published.

Table 2. Pairwise Θ values for the three Washington mountain goat populations and the south Cascades goats classified as ONP-Cascades admixed are given below the diagonal and the associated *P values* are given above the diagonal.

| | North Cascades | South Cascades | Olympic NP | ONP-Cascades admixed |
|---|---|---|---|---|
| North Cascades | - | 0.008 | 0.008 | 0.008 |
| South Cascades* | 0.113 | - | 0.008 | 0.008 |
| Olympic NP | 0.378 | 0.417 | - | 0.017 |
| ONP-Cascades admixed | 0.146 | 0.118 | 0.194 | - |

*The south Cascades population does not include ONP-Cascades admixed samples in this table.

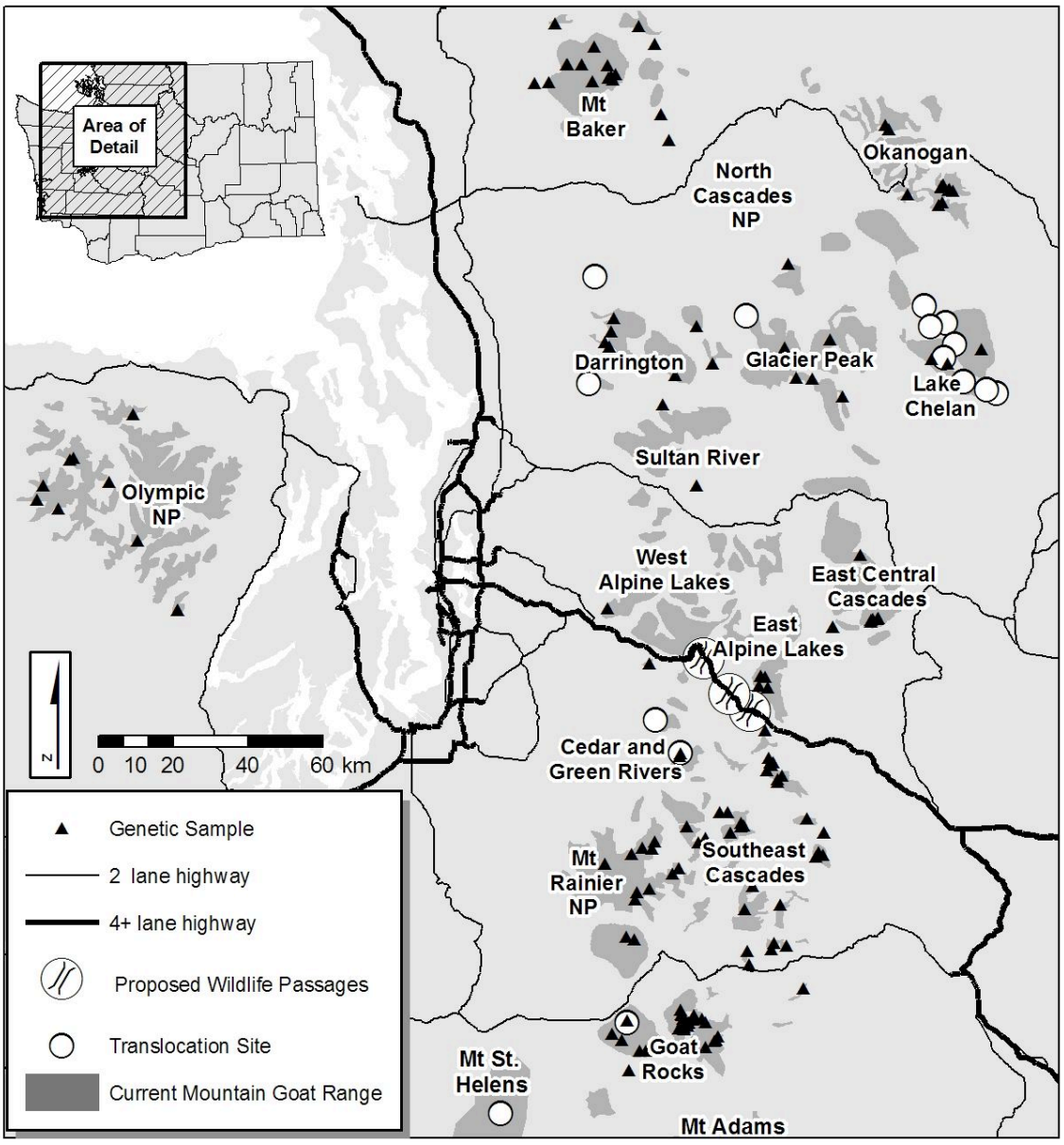


Figure 1. Map of the study area, including reference names to denote regions, location of genetic samples, highways, proposed highway passages, translocation sites, and the current mountain goat range.

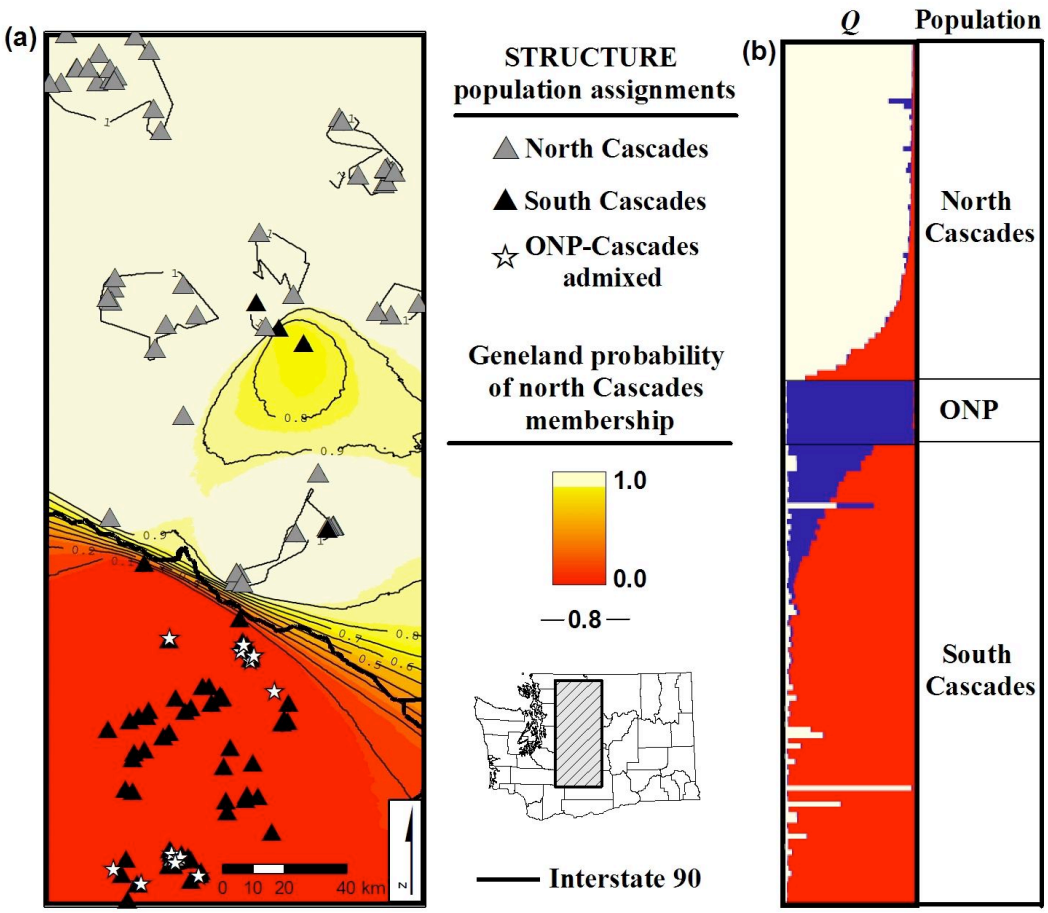


Figure 2. (a) Cascades sample locations are indicated by shapes that represent the population to which they have the highest probability of membership based on the STRUCTURE assignments. The Geneland posterior probability map of membership for the north Cascades is also represented. (b) STRUCTURE *Q*-Plot showing the posterior probability of membership (Q) for the south Cascades (red), north Cascades (cream), and ONP (blue). Each horizontal bar represents an individual, grouped by the population in which they were sampled (ordered by probability of north Cascades membership in the northern subpopulation and the probability of ONP membership in the southern subpopulation).

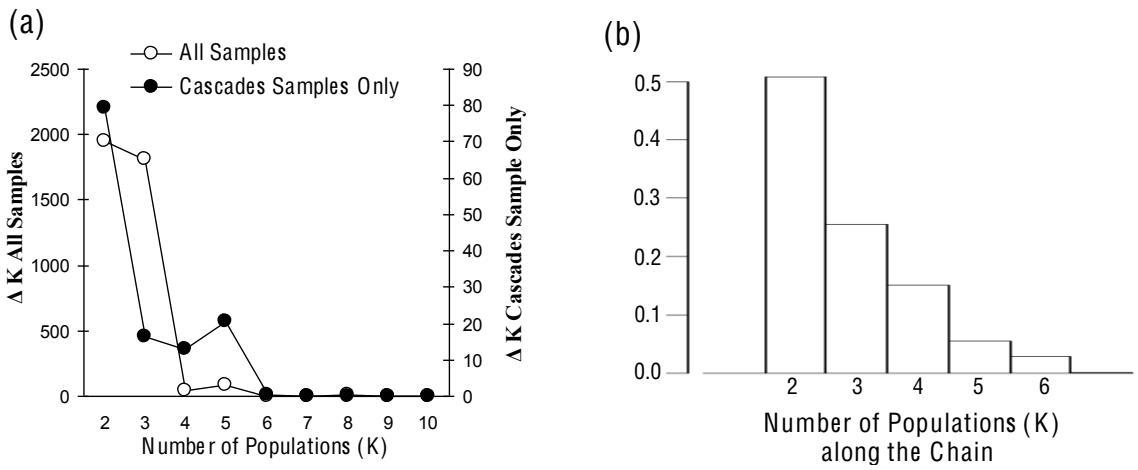


Figure 3. (a) The second order rate of change in the STRUCTURE assignment probability ($\Delta K$) for 1 to 10 populations. Open circles represent all samples and closed circles represent Cascade samples only (excluding ONP and ONP-Cascade admixed samples). (b) Density plot of the number of Cascade populations ($K$) along the Geneland MCMC chain.

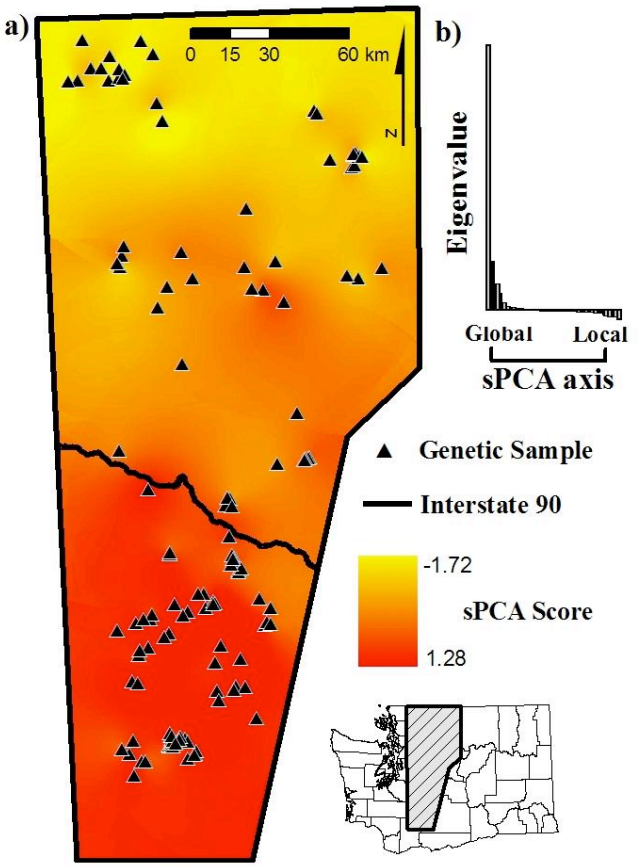


Figure 4. (a) Interpolated first global sPCA axis eigenvalues for all Cascade samples. (b) sPCA eigenvalues for each global and local axis.

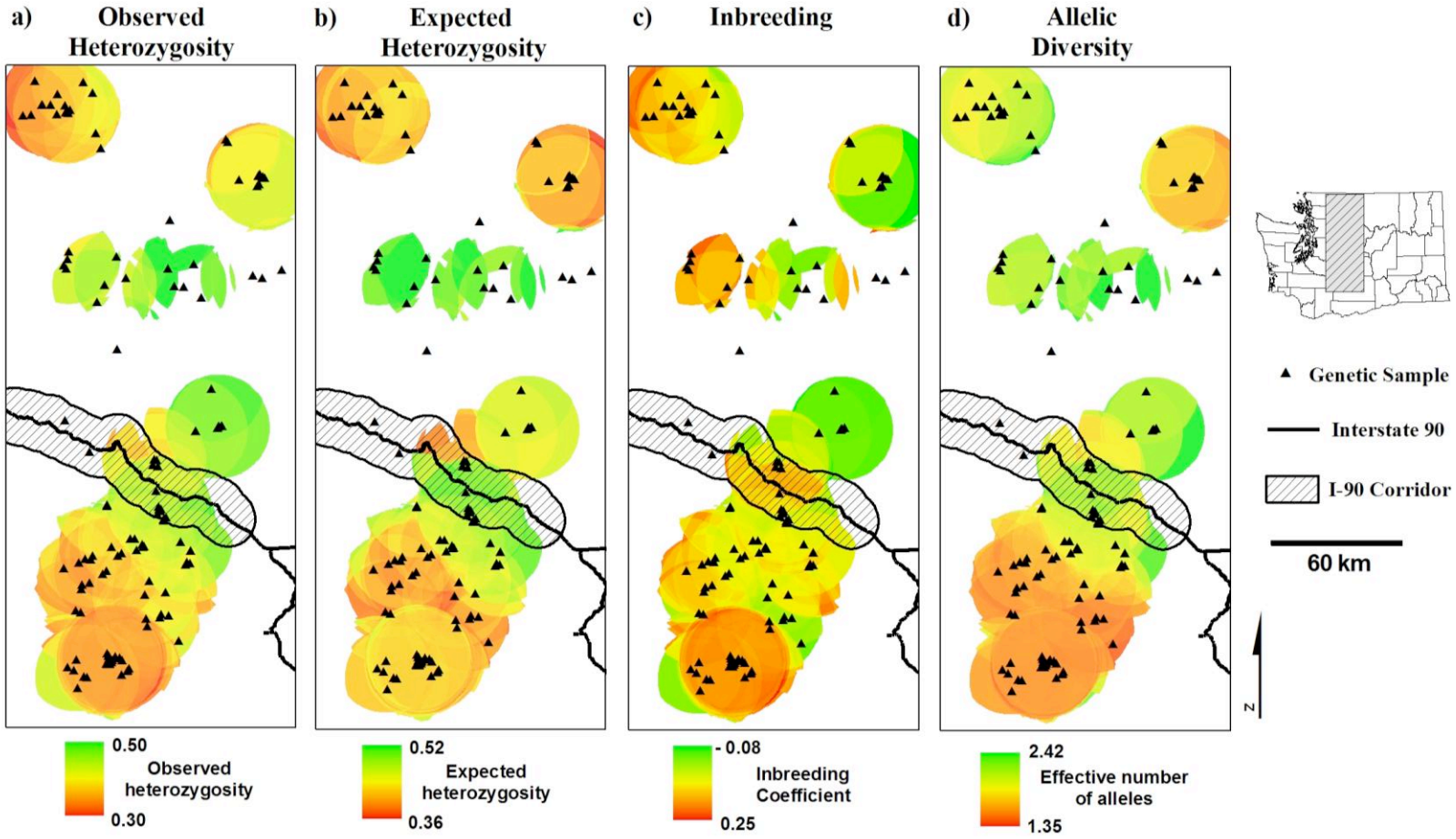


Figure 5. Genetic diversity indices calculated within a 24 km radius moving window, including (a) Observed Heterozygosity ($H_O$) (b) expected heterozygosity ($H_e$) (c) inbreeding coefficient ($F_{IS}$), and (d) effective number of alleles ($A_E$). Values that fall within the I-90 corridor region (crosshatched) may be inaccurate due to the barrier effect of the interstate.

# CHAPTER 2

## Mountain Goat Gene Flow in the Cascade Range, Washington

### INTRODUCTION

A primary effect of habitat fragmentation is to subdivide large populations into smaller demographically isolated subpopulations (Saunders et al. 1991). This, in turn, limits the flow of genetic variation across the landscape (gene flow) and reduces the local effective population size, $N_e$ (Keyghobadi 2007). $N_e$ is largely determined by the interaction between population size and gene flow, as well as demographic factors (e.g. sex ratios and mating success rates), and is a key factor governing the balance between accumulation of genetic diversity, through mutation and immigration, and loss of diversity through inbreeding and genetic drift (Frankham 1996). Small populations isolated from gene flow tend to reach an equilibrium that favors loss of genetic diversity, and are therefore susceptible to extinction due to genetic processes, including inbreeding depression (Crnokrak and Roff 1999), random fixation of deleterious alleles (Lynch et al. 1995, Lande 1998b), and loss of adaptive potential (Lande 1995, Swindell and Bouzat 2005). Improving connectivity in fragmented landscapes, and therefore the probability of population persistence, is now a major conservation priority (Crooks and Sanjayan 2006). Designing effective reserve systems and wildlife corridors for this purpose, however, requires an accurate understanding of the landscape factors which allow gene flow.

Gene flow is a function of the degree to which a landscape resists dispersal, which, in turn, is highly dependent on both the species and the landscape of interest (Hastings and Harrison 1994). Three conceptual models capture the ways in which this occurs. The

isolation by barrier (IBB) model predicts that strong barriers block dispersal, resulting in total isolation between subpopulations, but panmixia within each subpopulation. This mechanism of genetic isolation is commonly observed among species inhabiting habitat patches surrounded by an inhospitable matrix (e.g. aquatic species in unconnected lakes), and results in a discrete population structure.

For continuously distributed species, or where the matrix between patches is readily traversed, isolation by distance (IBD) is commonly observed (Wright 1943). IBD predicts that genetic isolation is a function of Euclidean distance, and results in populations structured along a distance gradient. For highly vagile species, genetic isolation may accrue slowly over very large distances, while sessile species often exhibit IBD over very short distances.

The third mechanism, isolation by landscape resistance (IBR), predicts that genetic isolation is a function of the degree to which a heterogeneous landscape resists migration (Cushman et al. 2006, McRae 2006). This model accounts for variation in the suitability of the matrix, and assigns a resistance cost to various landscape features (e.g. elevation, landcover type, roads, slope, or canopy cover) that affect migration. Under this model, genetic isolation accrues over distance as in IBD, but the rate at which it accrues varies according to the resistance of the landscape features. Isolation by resistance results in populations structured along a landscape resistance gradient, and is associated with species that inhabit complex landscapes that variably affect dispersal success.

Despite the importance of understanding the mechanism of genetic isolation to the conservation of fragmented populations, most studies do not evaluate the relative support for IBB, IBD, or IBR. Currently, the most common approach to identify population genetic structure is to assign individuals to discrete subpopulations based on assignment tests such as

STRUCTURE (Pritchard et al. 2000). Importantly, this approach will always assign individuals to one or more discrete subpopulations, even if the actual structure is a gradient arising from IBD or IBR. Incorrectly assuming discrete subpopulations isolated by barriers can lead to misinformed conclusions regarding genetic isolation and diminish the effectiveness of conservation efforts aimed at improving the viability of populations in fragmented landscapes.

In this study, I sought to determine the mechanism of genetic isolation in a population of mountain goats in the Cascade Range, Washington. This population has declined dramatically due to unsustainable levels of hunting over the past 50 years (Rice and Gay 2009), and inhabits a landscape that has been transformed by anthropogenic factors such as roads, timber harvest, and development. Although the number of hunting permits issued has been greatly reduced over the past 20 years, recovery among much of the remnant population has not been apparent (Rice, Washington Department of Fish and Wildlife, unpublished data). The potential for genetic factors to affect the recovery and long-term viability of this population is of management concern, and an understanding of genetic isolation is required to devise an effective conservation strategy.

In chapter 1, the discrete model based assignment tests STRUCTURE and Geneland suggested that the Cascades mountain goat population was divided into two subpopulations by an interstate (I-90), suggesting IBB was the predominant mechanism of genetic isolation within the study area (though some migration was observed between the north and south, indicating I-90 was not a complete barrier). However, spatial principal components analysis, sPCA (Jombart et al. 2008), revealed a north-south gradient of genetic isolation among mountain goats in the Cascades, suggesting IBD or IBR. To resolve this apparent

contradiction and develop a model of genetic isolation within the Cascades that can be used for mountain goat conservation, I sought to evaluate the relative support for each of these conceptual models.

Cushman et al. (2006) described a method to evaluate the relative support for IBB, IBD, and IBR within an individual-based causal modeling framework. This approach recognizes that, while IBB and IBD are easily modeled due to their simple assumptions, modeling IBR can be challenging due to the generally poor understanding of how a landscape resists gene flow for a particular species. While biologists typically have an understanding of which factors may be important in shaping gene flow, quantitatively assigning a resistance to these factors based on expert opinion is highly subjective. Instead, Cushman et al. (2006) generated a range of plausible hypotheses relating elevation, slope, roads, and landcover to black bear (*Ursus americanus*) gene flow. The full factorial combination of these hypotheses was then evaluated based on the Monte Carlo significance of the Mantel's correlation (Mantel 1967) between genetic distance (as measured by Bray-Curtis dissimilarity of multilocus microsatellite genotypes) and landscape resistance (quantified according to the total cost of the least cost-path connecting each pair of individuals) for each hypothesis. The relative support for IBR, IBD, and IBB was also evaluated with partial Mantel's tests (Smouse 1986). Using this approach, Cushman et al. (2006) found a model of IBR that included resistance due to landcover type, elevation, and slope was most strongly related to genetic isolation of black bears in northern Idaho.

In this study, I tested the support for IBB, IBD, and multiple hypotheses of IBR that varied in the resistance assigned to roads, landcover type, elevation, and distance to escape terrain, all factors I considered *a priori* to be important forces shaping mountain goat gene

flow. My approach was based on the Cushman et al. framework, however, I included a multivariate process of IBR model selection, used principal components analysis (PCA) to quantify genetic distance, and used circuit theory (McRae 2006) to model landscape resistance between sampling locations as well as gene flow. I conclude by describing the ways in which these modifications improved model selection, and discuss the implications of the most supported model of genetic isolation to the recovery and long-term viability of mountain goats in the Cascade Range, Washington.

## METHODS

### Study area

The study area in Chapter 2 represents a subset of the Cascade Range studied in Chapter 1, and includes the areas proximal to the Cascade genetic samples (Figure 1).

### Sample collection, and genotyping

For sample collection and genotyping, see chapter 1 methods. In Chapter 1, I identified 14 individuals among the 149 Cascade samples collected that were significantly related to the non-native Olympic National Park (ONP) population. These ONP-Cascade admixed goats are likely descendants of animals translocated from ONP to the Cascades in the 1980's. The location of their genetic variation, therefore, is highly influenced by management history, rather than the effect of the landscape per se. As such, they were excluded from the landscape genetics analysis of Chapter 2.

## GIS data

I used ArcGIS 9.2 (ESRI) to generate a slope raster based on a 30 m resolution digital elevation model (DEM) obtained from the USGS (USGS). I then used the ArcGIS "Euclidean Distance" function to generate a raster of Euclidean distance to escape terrain, $D_{et}$, (defined as slope $\geq 50°$ based on a habitat model derived for coastal mountain goats, (Smith 1994). I obtained landcover data (raster, 30 m cell size) from the Northwest GAP Analysis Project (NWGAP), and reclassified it according to 7 broad categories, including alpine/rock, subalpine forest/parkland, montane forest, mesic forest/disturbed, urban/ agricultural, ice, and water. I also obtained road data (ESRI shapefile) from the Washington State Department of Transportation (WSDOT). I converted this layer to a raster (30 m cell size) and reclassified it into four broad categories: no road, secondary roads, highway, and interstate. All GIS data were projected to UTM NAD27 Zone 10 North.

## Quantifying genetic distance

In molecular studies, genetic distance between individuals is commonly estimated with a model that assumes a particular mechanism of mutation, such as the infinite alleles model (Kimura and Crow 1964) or the stepwise mutation model (Ohta and Kimura 1973). As these models make a number of assumptions which may not be valid, particularly for highly variable markers like microsatellites (Balloux and Lugon-Moulin 2002), I chose to use principal components analysis (PCA) to estimate genetic distance. PCA generates synthetic variables (principle component axes) which reduce multidimensional data (in this case multilocus microsatellite genotypes) into a single dimension that contains most of the

variance (1st PCA axis) and makes no biological assumptions. I generated a data frame in R 2.8 (The Comprehensive R Archive Network, CRAN) containing all individuals in my dataset in rows and all alleles present in the Cascade population in columns. I then scored each allele as 0 (not present), 1 (heterozygous), or 2 (homozygous) for all individuals and used R 2.8 to calculate the 1st PCA axis eigenvalues. I then used the Ecodist package in R 2.8 (Goslee and Urban 2007) to generate an $n$ x $n$ pairwise genetic distance matrix based on the Euclidean distance between individuals along the 1st PCA axis ($n$ = 135).

### Modeling IBB and IBD

I modeled IBB by assigning samples ($n$ = 135) to a north or south subpopulation based on their sampling location relative to the hypothetical barrier (I-90). I then generated an $n$ x $n$ distance matrix where pairwise distances between samples within the same subpopulation were given a value of 0 and samples on opposite sides of the barrier were given a value of 1.

To model IBD, I used a Visual Basic script (written by Li Zou, available from ESRI) in ArcGIS 9.2 (ESRI) to generate an $n$ x $n$ matrix of Euclidean distance (in meters) between all sampling locations ($n$ = 135) based on their UTM coordinates. As the Cascade Range is essentially a linear north-south landscape, untransformed distances are likely to be most appropriate, however, I also tested the correlation of a $Log_{10}$ transformed IBD model, which would theoretically exhibit a stronger relationship with genetic distance in two-dimensional landscapes (Rousset 1997).

### Modeling IBR

I modeled IBR by generating one $n$ x $n$ matrix of resistance distances for each IBR resistance hypothesis I tested. I used Circuitscape 3.0.1 (McRae 2006) to populate each matrix with

pairwise resistance distances between all sample locations, given the resistance hypothesis. Circuitscape uses graph and circuit theory to calculate resistance between points given a GIS raster layer defining landscape resistance (i.e. a resistance surface). I used a connection scheme where gene flow was only allowed between the nearest 4 cells (i.e. no diagonal connections) and resistance between any two cells was based on the average of the resistance assigned to both cells. In order to achieve a reasonable computing time for each Circuitscape run (average about 2 hours), I lowered the resolution of each resistance surface from 30 m to 450 m by aggregating 15x15 blocks of 30 m pixels into a single pixel based on the average resistance of all the cells within the block.

### Evaluating the relative support for IBB, IBD, and IBR

I used the R software package Ecodist (Goslee and Urban 2007) to evaluate the support for multiple hypotheses of IBR, as well as IBB and IBD, based on the strength of the Mantel's correlation between each landscape distance matrix (barrier distance, Euclidean distance, or resistance distance) and the genetic distance matrix. I also used partial Mantel's tests to assess the relative support for each conceptual model (IBB, IBD, or the most supported IBR hypothesis) while controlling for the effect of one of the other models.

### Univariate IBR model selection

I evaluated the relative support for multiple hypotheses of IBR based on the following steps: 1) identify factors likely to contribute to landscape resistance 2) identify an expert opinion model (based on a thorough review of the available literature, consultation with experts, and data obtained from 39 GPS collared mountain goats in the Cascade Range) that assigns a

resistance cost to each of those factors 3) test the assumptions of the expert opinion model by evaluating alternative hypotheses of landscape resistance for each factor and determining if any of these alternatives improved the fit of the model to the data (univariate optimization), and 4) combine the univariate optimal hypothesis for each factor into a multivariate model, and evaluate variations of this model to determine if alternate multivariate hypotheses improve the fit of the model to the data (multivariate optimization, described in the next section).

To begin, I defined *a priori* four factors that I hypothesized were important resistors of gene flow among mountain goats, including distance to escape terrain, landcover type, roads, and elevation. For each of these factors, I generated an expert opinion hypothesis that assigned a resistance to every pixel in a GIS raster representing that factor. Because of the difficulty in reliably estimating the resistance of a landscape to gene flow based on expert opinion, I also generated additional hypotheses, as described below.

To generate an expert opinion model accounting for the resistance due to increasing distance to escape terrain ($D_{et}$), I reclassified a raster representing Euclidean distance to escape terrain (slope > 50) according to the following function:

Resistance (R) = $D_{et}^{x}$

where the response shape exponent ($x$) = 0.1 in the expert opinion hypothesis. I then linearly rescaled the result to range from 1 to 50 (Figure 2a), with the maximum resistance occurring at ≥ 600 m. This cutoff value was chosen based on data collected from 39 GPS collared mountain goats in the Cascade Range, Washington (Cliff Rice, WDFW, unpublished data). These data suggested that the frequency with which mountain goats are observed decreases exponentially with distance to escape terrain, and 95% of all observations were within 600 m.

Next, I generated 4 additional hypotheses about this expert opinion model that either increased or decreased the response shape exponent or the maximum scale of the response. If increasing or decreasing the response shape exponent/scale improved the fit of the model (evaluated as described below), I generated additional hypotheses further increasing/decreasing the shape exponent/scale until a unimodal peak of support (i.e. the fit of all other models tested increased steadily towards the optimal model) was achieved.

I assigned road resistance for each of four categories of road type (no road, secondary/forest service road, highway, and interstate), each of which was characterized by the annual average daily traffic volume (Washington State Department of Transportation, 2006) from July 2006 to October 2006 (0, 50, 4,000, and 28,000, respectively) according to the function:

$$R = \text{annual average daily traffic volume}^{x}$$

where $x = 1.5$ for the expert opinion model. I then rescaled the result to range from 1 (for non-road pixels) to 1000 (Figure 3a). I generated additional road resistance hypotheses by increasing/decreasing the shape exponent/scale to achieve a unimodal peak of support.

Because landcover type is categorical, I modeled its resistance by first ranking each of the seven landcover types in order of increasing resistance. This order was informed by a Cascade Range mountain goat habitat study (Wells 2006). I assigned alpine/rock the least resistance, followed by, in order of increasing resistance, subalpine forest/parkland, ice, montane forest, mesic forest/disturbed, and water. I assigned landcover resistance according to the function:

$$R = \text{landcover rank}^{x}$$

where $x = 5$ for the expert opinion model. I then rescaled the result to range from 1 to 50 (Figure 4a). In addition to the above landcover types, some urban and agricultural areas were present within the study area. I reasoned that, because these categories are only present in small patches within the Cascades, and mountain goats avoid human habitats, mountain goat gene flow would be directed around rather than through these areas. The high cost of these areas relative to other landcover types was also difficult to capture in a simple formula relating landcover to resistance. To avoid this issue, and force gene flow around these areas, I chose to assign urban/agricultural areas a prohibitively high resistance fixed at 100,000. I generated additional landcover resistance hypotheses by increasing/decreasing the shape exponent/scale to achieve a unimodal peak of support.

I assigned elevation resistance based on a Gaussian function defined an optimal elevations and a standard deviations (1400 m and 1000 m, respectively for the expert opinion model). I then rescaled the result to range from 1 to 5 (Figure 5a). I reasoned that mountain goat gene flow would never traverse the summits of Cascade stratovolcanoes and therefore assigned this elevation class (> 3,300 m) a prohibitively high resistance fixed at 100,000. No mountain goats have been observed (by GPS collar data or visual sighting) above this elevation in the Cascade Range (Rice, WDFW, personal communication). I generated additional elevation resistance hypotheses by increasing/decreasing the shape exponent/scale to achieve a unimodal peak of support.

### Multivariate IBR model selection

Once I had determined an optimal univariate model for each of the four factors considered, I then combined these factors into a multivariate model. To eliminate factors that might have been spuriously related to landscape resistance due to the number of hypotheses tested, I set the following criteria for inclusion of a factor in the multivariate model: 1) the Mantel's correlation was greater than the null model of isolation by distance, in which all cells have a resistance of one 2) the Mantel's correlation represents a unimodal peak of support (the Mantel's correlation for all other hypotheses for a given factor increases towards the optimal model), and 3) The *P*-value was significant after Bonferroni correction for multiple tests (Rice 1989).

I then used raster algebra in ArcGIS 9.2 (ESRI) to additively combine the univariate resistance hypotheses that met the above criteria into a single multivariate model. The optimal shape exponent/scale for each factor could potentially change in a multivariate context. Optimizing the shape exponent/scale of all factors simultaneously was not practical due to the sheer number of potential combinations. Instead, I chose to vary the shape exponent and scale of one factor while holding the others constant. In some cases, this involved extending the range of the shape exponent or scale beyond what was used in the univariate optimization so that a clear peak of support was achieved. I optimized the factors in order of decreasing Mantel's correlation for the univariate optimal hypotheses. Once the first factor was optimized in the multivariate context, I held it constant and varied the shape/scale of the next most highly correlated factor. If the shape exponent/scale of any factor was changed in the multivariate context, I repeated the optimization process in a second round.

### Predicting gene flow

Given a resistance surface hypothesis, it is possible to use circuit theory to calculate the flow of current (analogous to gene flow) between two or more points across that surface. To illustrate this, I used the most highly supported multivariate IBR hypothesis raster as the basis to calculate current between Mount Baker, in the north Cascades, and Mount Adams, in the south Cascades (points 1 and 4, Figure 10b). I used Circuitscape 3.1 (McRae 2006) to calculate current for all cells based on applying a 1 amp current to the Mount Adams point and connecting the Mount Baker point to ground. Current represents the relative frequency with which individuals moving by a "random walk" pattern would pass through a given cell as it travels across the resistance surface from the source to the ground point (McRae et al. 2008).

## RESULTS

### Univariate optimization

The $D_{et}$ hypothesis most correlated with genetic distance ($r = 0.669$, $P < 0.0001$, Figure 2b) was scaled from a resistance ranging from 1 to 5 and had a shape exponent of 0 (i.e., the highest resistance was achieved at any $D_{et} > 0$). Support for this hypothesis reached a unimodal peak, however, its correlation was lower than the null model ($r = 0.682$), and I therefore chose to eliminate $D_{et}$ as a factor in the multivariate model.

The roads resistance hypothesis most correlated with genetic distance ($r = 0.707$, $P < 0.0001$, Figure 3b) was scaled from a resistance of 1 to 2,500 and had a shape exponent of $\infty$, which assigns a resistance of 1 to all non-road, secondary road, and highway pixels and 2500

to interstate pixels. This hypothesis met all three criteria for inclusion in the multivariate model.

The landcover resistance hypothesis most correlated with genetic distance ($r = 0.692$, $P < 0.0001$, Figure 4b) was scaled from a resistance of 1 to 10 and had a shape exponent of 7, which assigns a resistance of 2 to mesic/disturbed forest and 10 to water (all other pixels have the minimum resistance of 1). This hypothesis met all three criteria for inclusion in the multivariate model.

The elevation resistance hypothesis most correlated with genetic distance ($r = 0.686$, $P < 0.0001$, Figure 5b) was scaled from a resistance of 1 to 5 and its shape was defined by the Gaussian optimal elevation of 1600 m with a standard deviation of 1500 m. This model met all three criteria for inclusion in the multivariate model.

### Multivariate optimization

Because the optimal road resistance model displayed the highest correlation with genetic distance of any of the univariate models, I began the multivariate optimization with this factor. In the context of the univariate optimal elevation and landcover models, the optimal road resistance shape remained at ∞, however, the scale shifted from a maximum of 2500 to a maximum of 10,000 ($r = 0.719$, $P < 0.0001$, Figure 6). Next, I optimized landcover resistance, holding the optimal univariate elevation model and the new multivariate-optimized road resistance model constant. In this context, the optimal landcover resistance shape exponent shifted from a value of 7 to ∞, and the scale shifted from a maximum of 10 to a maximum of 25 ($r = 0.720$, $P < 0.0001$, Figure 7). Holding the multivariate optimized road and landcover resistance models constant, I varied the shape and scale of elevation

resistance. The optimal shape exponent and scale, however, did not change from the univariate optimum elevation ($r = 0.721$, $P < 0.0001$, Figure 8). Lastly, I performed an additional multivariate optimization to determine if the change in landcover resistance altered the optimal shape/scale of road resistance. While the optimal road resistance scale did not change, the shape shifted from an exponent of $\infty$ to 3, which assigned the minimal resistance (1) to non-road and secondary road pixels, 80 to highways, and 10,000 to interstates ($r = 0.723$, $P < 0.0001$, Figure 9). The resistance costs associated with the final multivariate-optimized IBR model is given in Table 1 and a map depicting these resistances within the study area is given in Figure 10a.

### Relative Support for IBB, IBD, IBR

The IBB model displayed a strong correlation with genetic distance ($r = 0.645$, $P < 0.0001$, Table 2). The IBD model appeared more supported compared to the $IBD_{LOG10}$ model ($r = 0.686$, $P < 0.0001$ and $r = 0.538$, $P < 0.0001$, respectively, Table 2), and the multivariate optimized IBR model appeared more supported than the expert opinion IBR model ($r = 0.723$, $P < 0.0001$ and $r = 0.640$, $P < 0.0001$, respectively, Table 2). I used partial Mantel's tests to evaluate the relative support for the three conceptual models of genetic isolation (IBB, IBD, and the multivariate optimized IBR model). The IBD model retained significant relationship with genetic distance when the variance due to IBB was partialled out ($r = 0.419$, $P < 0.0001$, Table 2), but not when IBR was partialled out ($r = -0.006$, $P = 0.5916$ Table 2). The IBR model retained a significant relationship with genetic distance when the variance due to IBB or IBD was partialled out ($r = 0.426$, $P < 0.0001$; $r = 0.302$, $P < 0.0001$, respectively, Table 2). The IBB model retained a significant relationship with genetic

distance when the variance due to IBR or IBD was partialled out ($r = 0.076$, $P = 0.0007$ and $r = 0.300$, $P < 0.0001$, respectively, Table 2). Because the IBD and IBB models retained little or no correlation with genetic distance when the variance due to the IBR model was partialled out, while the converse was not true, the multivariate optimized IBR model appears to have the strongest relationship with genetic distance among the models tested.

### North-south landscape resistance and gene flow

Based on circuit theory, the north-south gene flow across the modern landscape reveals several pinch points where current is channelized (Figure 10b). As gene flow originating at Mount Adams (point 4, Figure 10b) moves northwards towards Mount Baker (point 1), more dispersal avenues become available and current travels in a broad swath through the south Cascades. As gene flow crosses I-90 and State Highway 2, a few developed areas and lakes/reservoirs constrain gene flow, resulting in several "pinch points" where current is increased in narrow channels. Further north, many available paths allow for a broad pattern of gene flow. Finally, as gene flow converges on Mount Baker, the urban and agricultural areas of the Skagit Valley force current to the east. As gene flow approaches the destination, current increases because all paths converge on a single point. The total resistance between Mount Baker and Mount Adams across the IBR resistance surface is 25.8. By comparison, the resistance between two points flanking I-90 (2 and 3, Figure 10b) is 12.5. The north-south resistance in the north Cascades (between points 1 and 2) is 14.2 while the north-south resistance in the south Cascades (between points 3 and 4) is 9.5.

## DISCUSSION

This study improves upon previous landscape genetics applications of causal modeling in several ways. Multivariate optimization of both the response shape and the resistance cost range of each IBR model factor avoids the arbitrary nature of assigning resistance based on expert opinion alone. This advantage was apparent in the comparatively poor performance of the expert opinion model relative to the optimized IBR model (Figures 2-5, Table 2), demonstrating that *a priori* assumptions regarding the shape and scale of landscape resistance may not be fully supported. Moreover, as the optimal univariate model changed when in the context of the other factors (Figures 6-9), multivariate optimization appears to be a worthwhile procedure to achieve the best possible fit of a resistance hypothesis to the data.

Although this approach to evaluating multiple hypotheses of landscape resistance allows for the assumptions of an expert opinion model to be explicitly tested, one limitation is the potential for spurious correlations to arise among the 120 IBR models I generated. The fact that all hypotheses for a given factor were ordinated (both by the response shape exponent and the resistance scale) allowed validation by the expectation of a unimodal peak of support for a single hypothesis per factor. Excluding factors that did not meet this criterion greatly minimized the potential for spurious correlations to influence model selection.

In addition, by excluding factors that do not exhibit a stronger correlation with genetic isolation than the null model of isolation by distance (all pixels are assigned a resistance of 1), the model selection process I used represents a rigorous and reproducible approach to identify important factors among several possible factors shaping gene flow. Moreover, this multivariate optimization process allows for the scale of each factor to shift independently of other factors, resulting in a natural weighting of those factors that most

strongly shape gene flow. In the optimized IBR model, for example, the scale of road resistance varied from 1 to 10,000 (depending on the road type), while the scale of elevation resistance ranged from 1 to 5 (depending on the elevation above or below the optimal elevation of 1600 m). This disparity reflects the relatively greater effect roads have on resistance, per pixel.

In another departure from previous causal modeling studies of IBR, I used circuit theory to quantify landscape resistance. More commonly, least-cost-path (LCP) analysis has been employed to quantify landscape resistance. Important limitations to LCP models are the unfounded assumptions that dispersing animals have perfect knowledge of the landscape, only use a single corridor, and encounter equal resistance whether the corridor is wide or constrained. By contrast, circuit theory assumes a random walk dispersal behavior, simultaneously integrates the contribution of all possible pathways to gene flow, and attributes greater resistance to narrow corridors than wide corridors (McRae 2006). This more robust theoretical foundation has been attributed to the ability of circuit theory to outperform LCP models in both simulated and real landscapes (McRae 2006, McRae and Beier 2007). In addition, circuit theory can also be used to calculate current, given a resistance surface. As current is a function of movement probability along all possible branches of a graph linking two nodes (McRae et al. 2008), it can be used to identify pinch points where gene flow is constrained. Such areas are natural targets for mitigation efforts aimed at improving gene flow if connectivity is found to be insufficient.

Another aspect of this study that potentially improves model selection is the use of PCA to quantify genetic distance. With this approach, alleles that capture the greatest proportion of the genetic variation in a population have a correspondingly greater

contribution to genetic distance than common alleles, which capture relatively little of a population's genetic variation. In theory, this makes PCA a more sensitive method to detect genetic dissimilarity than other methods, such as the Bray-Curtis dissimilarity index used in Cushman et al. (2006), where all alleles contribute equally. In addition, as PCA makes no biological assumptions, this approach is more parsimonious than distance metrics that assume an infinite alleles model (Kimura and Crow 1964) or a stepwise mutation model (Ohta and Kimura 1973) of microsatellite variation. While these theoretical advantages of PCA-based genetic distance are compelling, the performance of this method relative to other metrics of has not been tested.

Despite the relatively small sample size available for this study, I was able to derive a model of genetic isolation that exhibited among the strongest correlations with genetic distance yet published. This is likely attributable to the focus here on the individual as the unit of observation. In classical population genetics, the population is the unit of observation. This approach first requires assigning sampled individuals to discrete subpopulations, which is problematic for continuously distributed species. Furthermore, population-based metrics of genetic distance such as linearized $F_{ST}$ typically assume an island model of population structure (Slatkin 1995), which may not be representative of the true structure. Reliably estimating population-based metrics of genetic distance also requires a sample size on the order of tens of individuals, limiting the number of populations that can practically be delineated. This, in turn, limits the number of observations which can be used to correlate genetic distance with models of genetic isolation. The individual-based approach used here obviates the need for assigning individuals to a population and dramatically increases the

number of observations with which to make correlations. This is particularly advantageous for rare or difficult to sample species such as mountain goats.

Employing individual-based causal modeling, a multivariate model selection process, circuit theory, and PCA, I derived a model of genetic isolation by landscape resistance that was strongly correlated with genetic distance and was better supported than alternative models of genetic isolation (IBB, IBD) or an expert opinion IBR model (Table 2). The optimized IBR model broadly agrees with the findings presented in Chapter 1 in that I-90 is the most important single factor shaping genetic structure in the population. The greater support for this model of IBR relative to IBB, however, demonstrates that the resistance due to I-90 only captures a portion of the genetic isolation present in this landscape. In fact, there appears to be as much genetic isolation within the north and south subpopulations as there is between them (Figure 10b). Thus, the common approach of assuming panmixia within discrete subpopulations isolated by barriers, in this case, would grossly overestimate gene flow within the north and south Cascades. Similarly, if I had only evaluated the support for IBD (as is commonly performed in population genetic studies) I would have overstated gene flow both within each subpopulation and across I-90 by not accounting for the isolation arising from roads, elevation, and the resistive landcover types in the optimized IBR model. The fact that all three models of genetic isolation exhibited strong and significant Mantel's correlations with genetic distance highlights the need to rigorously evaluate the relative support for IBB, IBD, and IBR.

While the most supported IBR model in this study largely fit the expectations of genetic isolation suggested by the Chapter 1 analysis, there were two surprising findings. First, the most supported response shape for landcover assigned the same low cost (the

minimum resistance of 1) to both forests and alpine terrain. While mountain goats are typically associated with alpine terrain in summer months when low snowpack permits dispersal, the coastal North American mountain goat populations, unlike interior populations, frequently move to lower elevations below treeline in winter (Demarchi et al. 2000, Festa-Bianchet and Cote 2007). As such, Cascade mountain goats may be more accustomed to these environments and not be strongly inhibited from moving through forests. The potential for this adaptation is partially supported by GPS data collected from Cascade mountain goats, which reveals several long-distance dispersal events and movements to mineral licks that involved traversing large tracts of continuous forest (Rice, Washington Department of Fish and Wildlife, unpublished data). Alternatively, because landcover type and elevation are correlated, the cost of forests may be at least partially encompassed by the resistance due to elevation included in the model.

Another unexpected finding was the exclusion of resistance due to escape terrain from the most supported IBR model. This species is known to rely heavily on escape terrain for predator avoidance, and, as a result, mountain goat movements appear highly correlated with proximity to steep slopes (Gross et al. 2002, Festa-Bianchet and Cote 2007). A possible behavioral explanation for this finding may be that dispersing goats in unfamiliar terrain do not have the perceptual range to detect routes that maximize escape terrain. Alternatively, a methodological explanation may stem from the 450 m resolution I used to model landscape resistance, which may not closely match the scale at which escape terrain is represented in the landscape, perhaps masking the importance of this factor. This coarser resolution was necessary due to the impractical computing time required for a finer scale analyses.

The gene flow predicted by the IBR model has important implications for the viability of mountain goats in the Cascade Range. Urban and agricultural areas were categorized as highly resistive landscape features in this model. Consequently, the Puget Sound lowlands would be predicted to force gene flow eastward. The effect of this altered gene flow would be predicted to increase the genetic isolation between mountain goats inhabiting the western extremes of the Cascades, including groups in the Mount Baker, Darrington, and western Alpine Lakes region. The few developed areas within the core of the Cascade Range, however, do not appear to generate pinch points where north-south gene flow is greatly constrained, except between a few urban areas along State Highway 2 and I-90. In contrast, no path linking the north and south Cascades avoids the high resistance of I-90.

As the local effective population size governs the balance between gain and loss of genetic diversity, the reduction in $N_e$ due to the high resistance of I90 would be predicted to favor loss of genetic diversity over time. Reduced genetic diversity, in turn, would be predicted to increase the risk of inbreeding depression in the short term as well as loss of adaptive potential fixation of deleterious alleles in the long term, thus negatively impacting the population's future viability. The other factors predicted to reduce gene flow in the IBR model, including suboptimal elevations, highways, and certain landcover types, though they appear to have a relatively modest resistance per pixel, have a cumulative effect comparable to that of I-90 (Figure 10b), resulting in a further reduction in effective population size and additional extinction risk to the population. These effects of genetic isolation due, in part, to modern landscape changes suggest the need to estimate $N_e$ considering the importance of this parameter to population viability. The understanding of factors affecting landscape resistance

gained in this study may be used to quantify $N_e$ given the resistance of the modern landscape. This may then be used as a means to quantify the potential for genetic factors to influence the recovery and long-term viability of mountain goats in the Cascade Range under a range of scenarios, including the construction of I-90 wildlife passages and other potential landscape changes.

# TABLES

Table 1. Resistance cost for each landscape factor contributing to resistance in the most supported isolation by resistance model.

| Factor | Resistance | |
|---|---|---|
| Urban/Agriculture | 100,000 | Fixed |
| Elevation > 3,300 m | 100,000 | Fixed |
| Interstate 90 | 10,000 | Optimized |
| Highways | 80 | Optimized |
| Water | 25 | Optimized |
| Elevation < or > 1600 m | 2 - 5 | Optimized |
| All other pixels | 1 | Optimized |

Table 2. Mantel's correlation with genetic distance (G) for each model of genetic isolation and partial Mantel's correlations with genetic distance, controlling for alternative models of genetic isolation with *P* values based on Monte Carlo randomization of the row/column order.

| Model | **Mantel's** *r* | Monte Carlo *P* value |
|---|---|---|
| IBB ~ G | 0.645 | < 0.0001 |
| IBD ~ G | 0.686 | < 0.0001 |
| $IBD_{LOG10}$ ~ G | 0.538 | 0.0001 |
| IBR ~ G | 0.723 | < 0.0001 |
| Expert IBR ~ | 0.640 | < 0.0001 |
| IBB ~ G \| IBD | 0.300 | < 0.0001 |
| IBD ~ G \| IBB | 0.419 | < 0.0001 |
| IBB ~ G \| IBR | 0.076 | 0.0007 |
| IBR ~ G \| IBB | 0.426 | < 0.0001 |
| IBD ~ G \| IBR | - 0.006 | 0.5916 |
| IBR ~ G \| IBD | 0.303 | < 0.0001 |

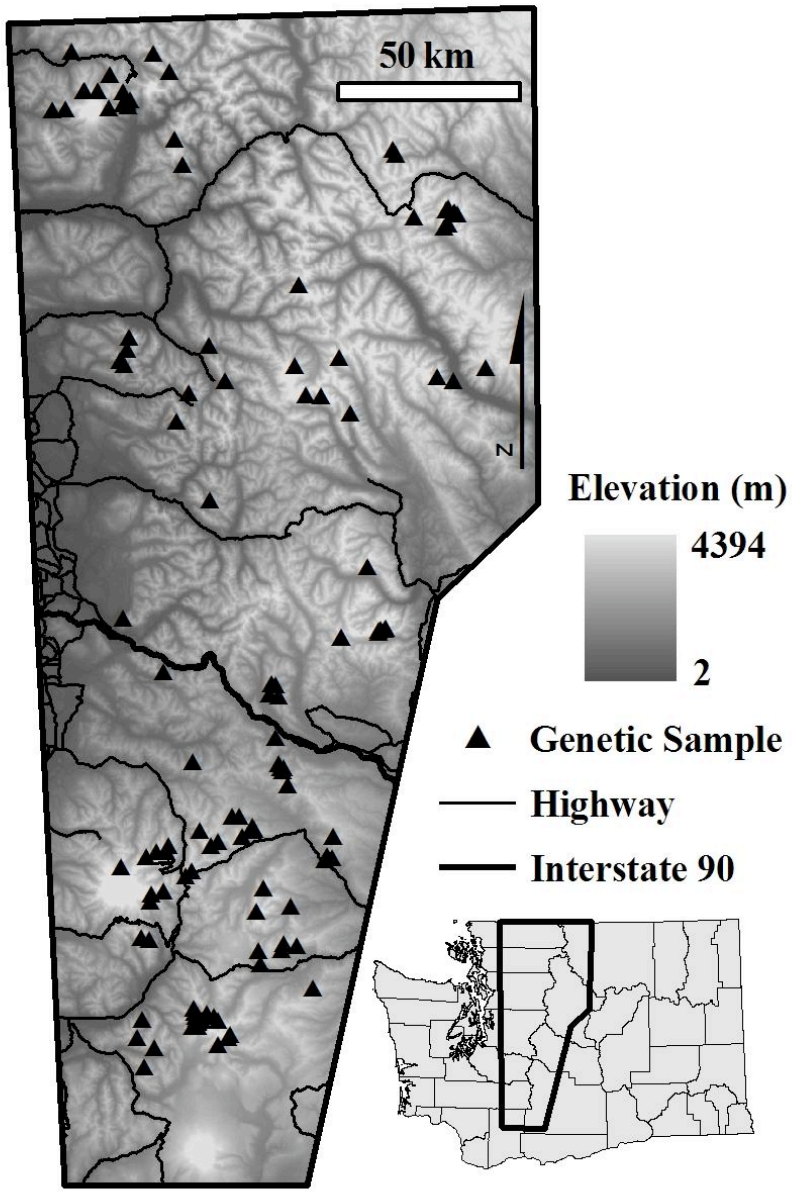


Figure 1. The study area is depicted as a digital elevation model in meters, with highways shown in thin black lines, Interstate 90 as a thick black line, and genetic samples denoted by black triangles.

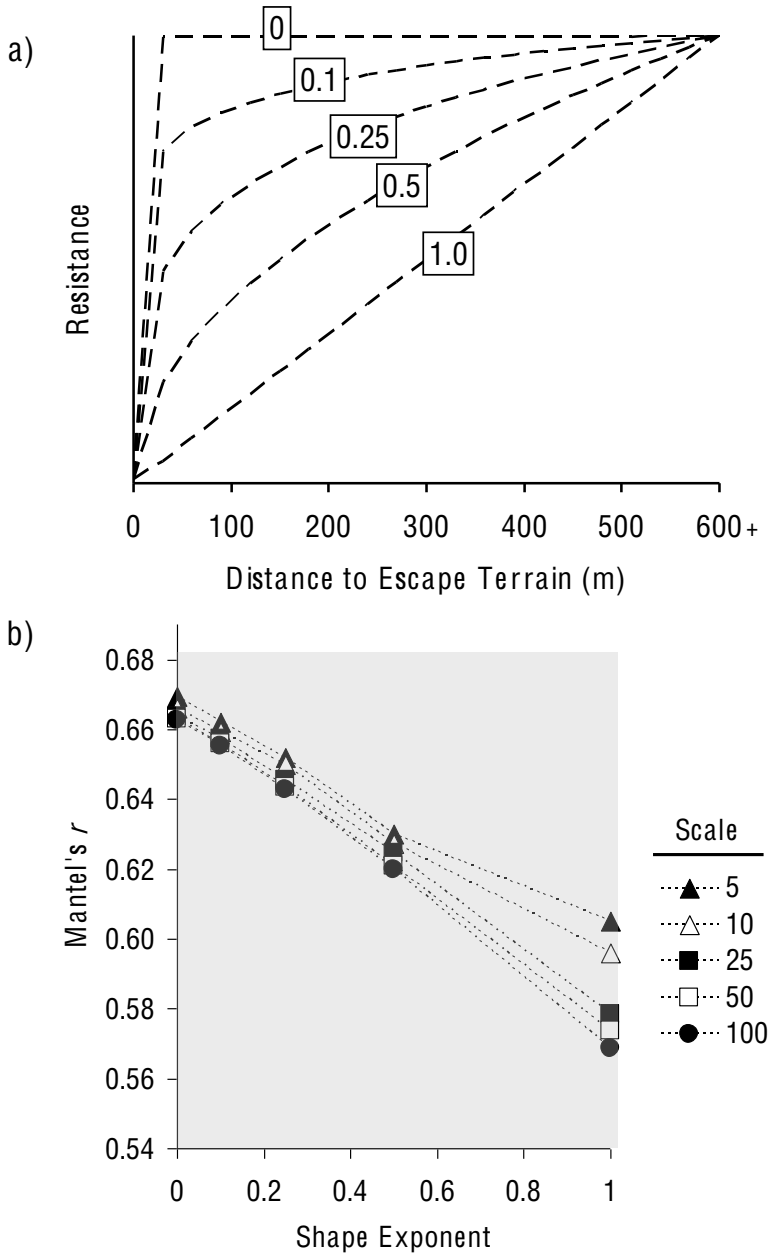


Figure 2. (a) The resistance assigned to increasing distance from escape terrain (resistance = distance $^{X}$, rescaled to range from 1 to the maximum resistance specified by the resistance hypothesis), given one of the five response shapes (the exponent, $X$, is noted in a box along each shape curve). (b) The Mantel's correlation with genetic distance is given for 25 hypotheses of resistance due to escape terrain that vary by $X$ (given on the x-axis) and the maximum resistance (given in the scale legend). The gray area indicates hypotheses less correlated with genetic distance than the null model. All hypotheses presented were significant ($\alpha = 0.05$) after Bonferroni correction for multiple tests.

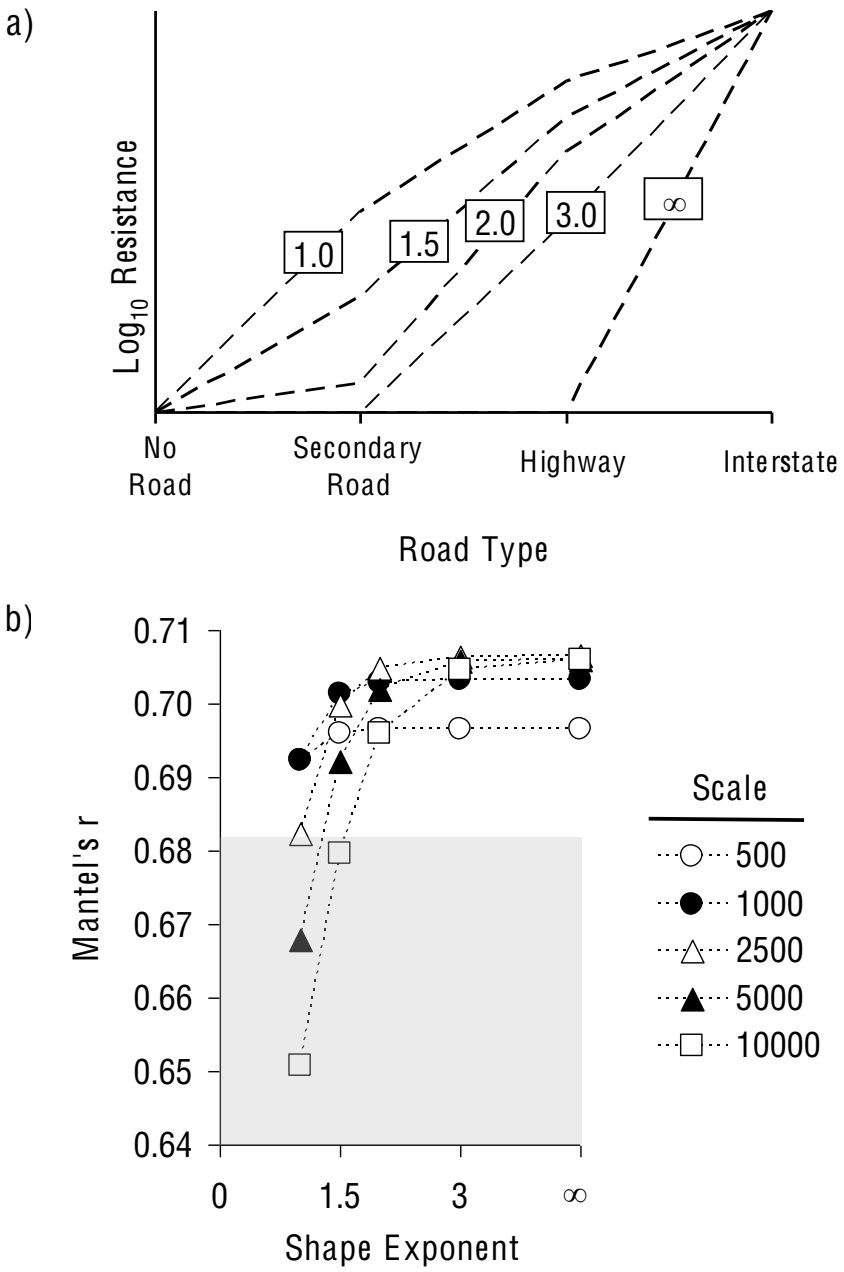


Figure 3. (a) The $\log_{10}$ of resistance assigned to various road types (resistance = traffic volume $^{\chi}$, rescaled to range from 1 to the maximum resistance specified by the resistance hypothesis), given one of the five response shapes (the exponent, $\chi$, is noted in a box along each shape curve). (b) The Mantel's correlation with genetic distance is given for 25 hypotheses of resistance due to roads that vary by $\chi$ (given on the x-axis) and the maximum resistance (given in the scale legend). The gray area indicates hypotheses less correlated with genetic distance than the null model. All hypotheses presented were significant ($\alpha = 0.05$) after Bonferroni correction for multiple tests.

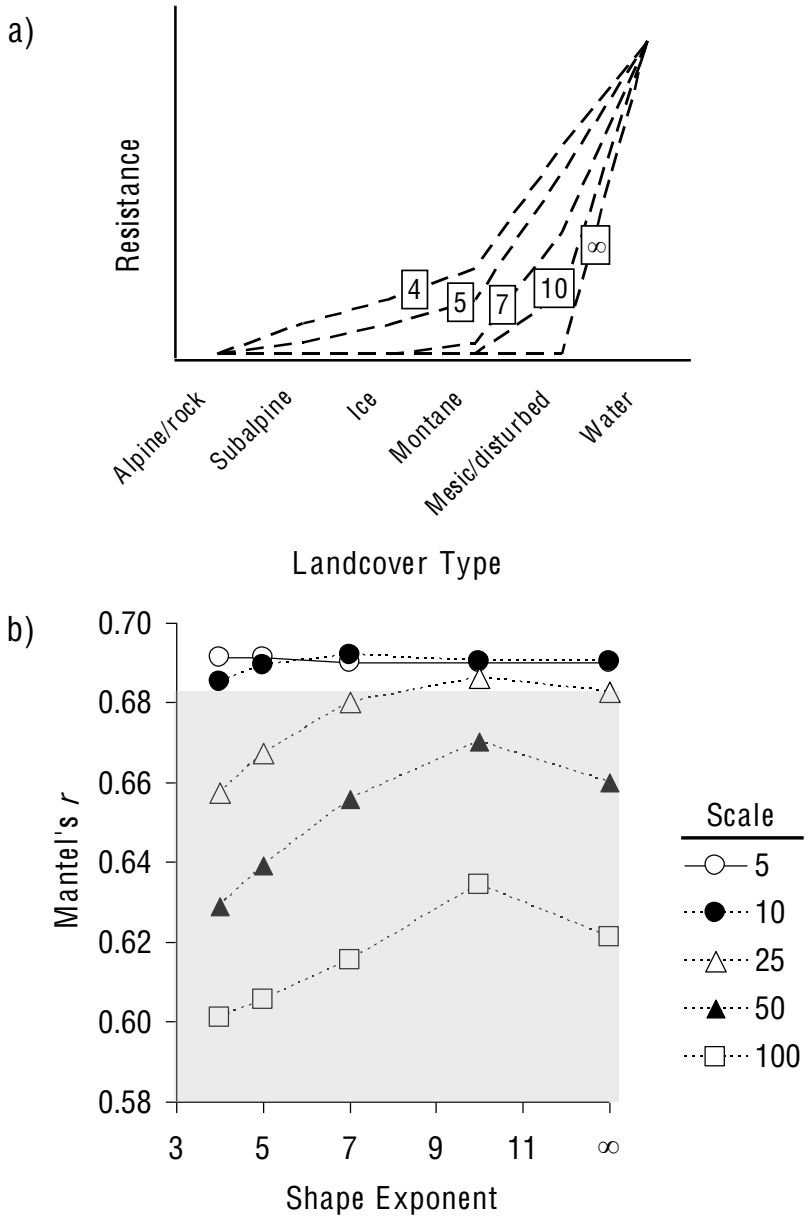


Figure 4. (a) The resistance assigned to various landcover types (resistance = landcover rank $^{\chi}$, rescaled to range from 1 to the maximum resistance specified by the resistance hypothesis), given one of the five response shapes (the exponent, $\chi$, is noted in a box along each shape curve). (b) The Mantel's correlation with genetic distance is given for 25 hypotheses of resistance due to landcover that vary by $\chi$ (given on the x-axis) and the maximum resistance (given in the scale legend). The gray area indicates hypotheses less correlated with genetic distance than the null model. All hypotheses presented were significant ($\alpha = 0.05$) after Bonferroni correction for multiple tests.

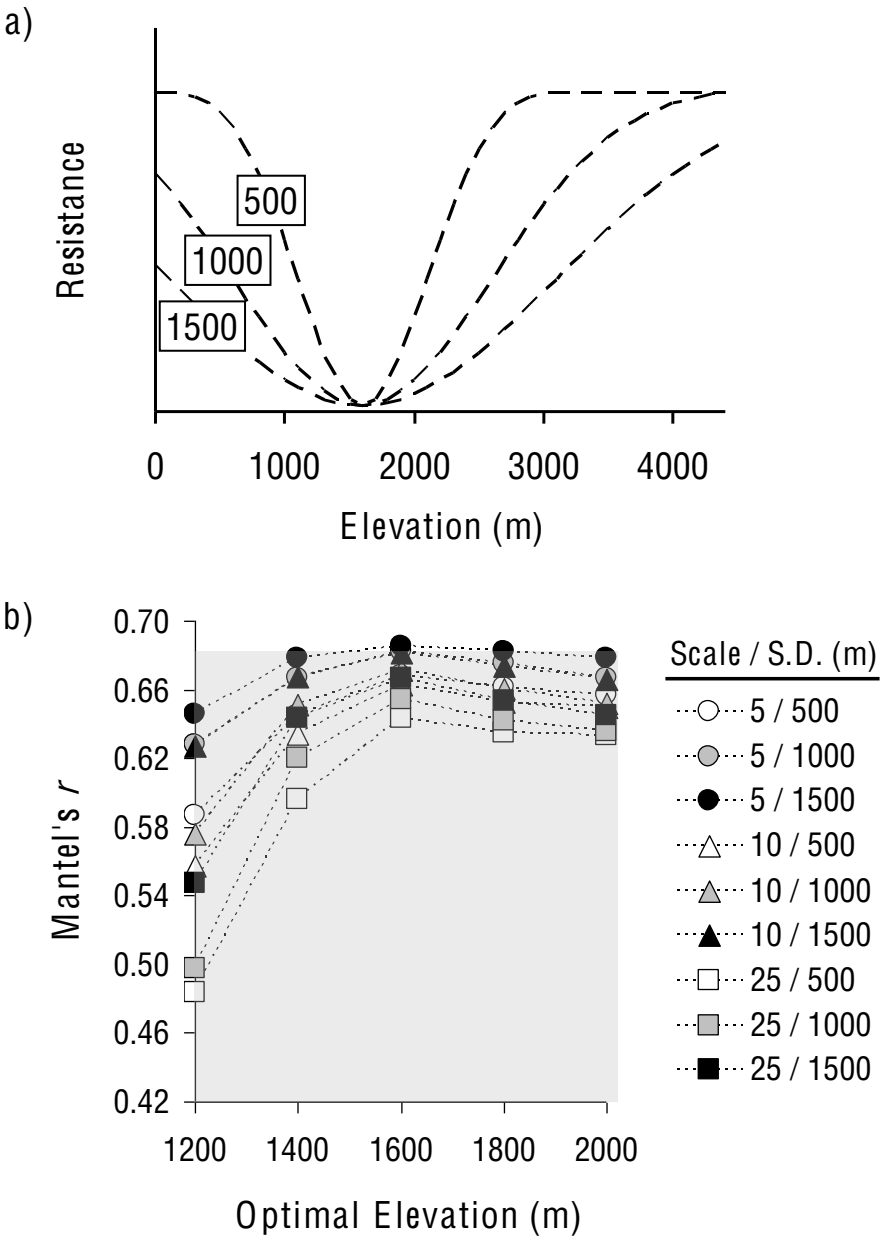


Figure 5. (a) The resistance assigned to various elevations (resistance = elevation $^{X}$, rescaled to range from 1 to the maximum resistance specified by the resistance hypothesis), given one of the five response shapes (the exponent, $X$, is noted in a box along each shape curve). (b) The Mantel's correlation with genetic distance is given for 25 hypotheses of resistance due to elevation that vary by $X$ (given on the x-axis) and the maximum resistance (given in the scale legend). The gray area indicates hypotheses less correlated with genetic distance than the null model. All hypotheses presented were significant ($\alpha = 0.05$) after Bonferroni correction for multiple tests.

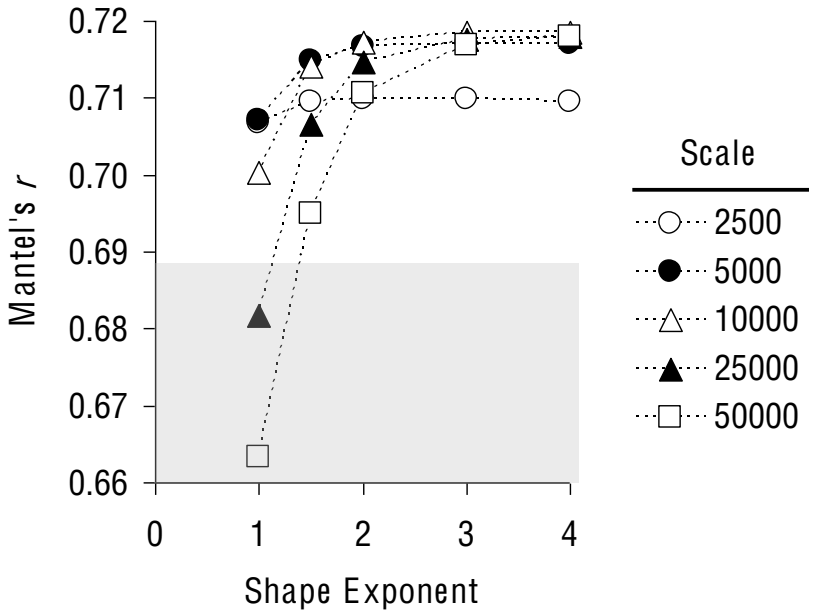


Figure 6. The Mantel's correlation with genetic distance is given for 25 hypotheses of resistance due to roads that vary by the shape exponent (resistance = road type $^{X}$) along the x-axis, rescaled to range from 1 to the maximum cost given in the scale legend. The resistance for each of these hypotheses was added to the most highly correlated univariate model of elevation and landcover resistance prior to testing the Mantel's correlation.

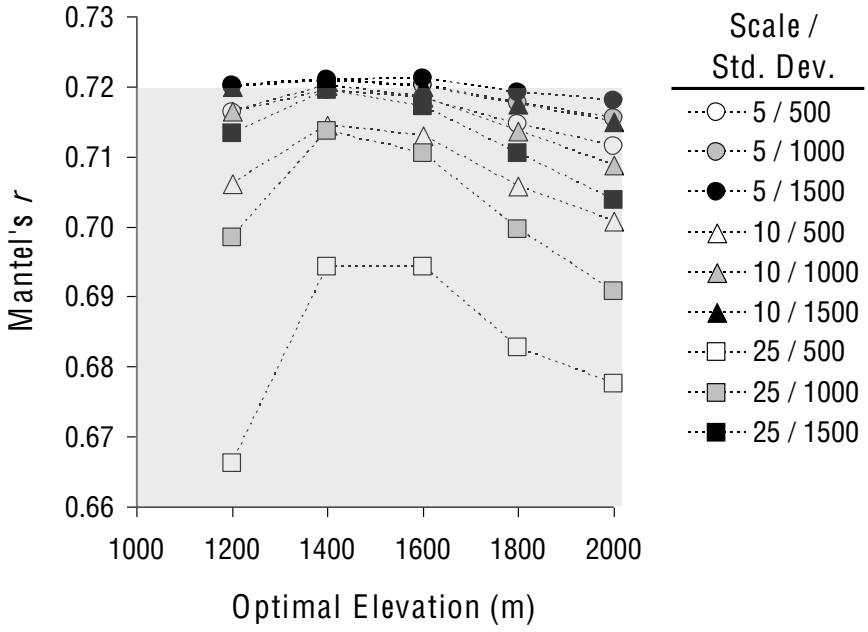


Figure 7. The Mantel's correlation with genetic distance is given for 45 hypotheses of resistance due to elevation that vary according to a Gaussian function defined by the optimal elevation along the x-axis and the standard deviation of the Gaussian elevation, rescaled to range from 1 to the maximum cost (see legend). The resistance for each of these hypotheses was added to the most highly correlated univariate model of road and landcover resistance prior to testing the Mantel's correlation.

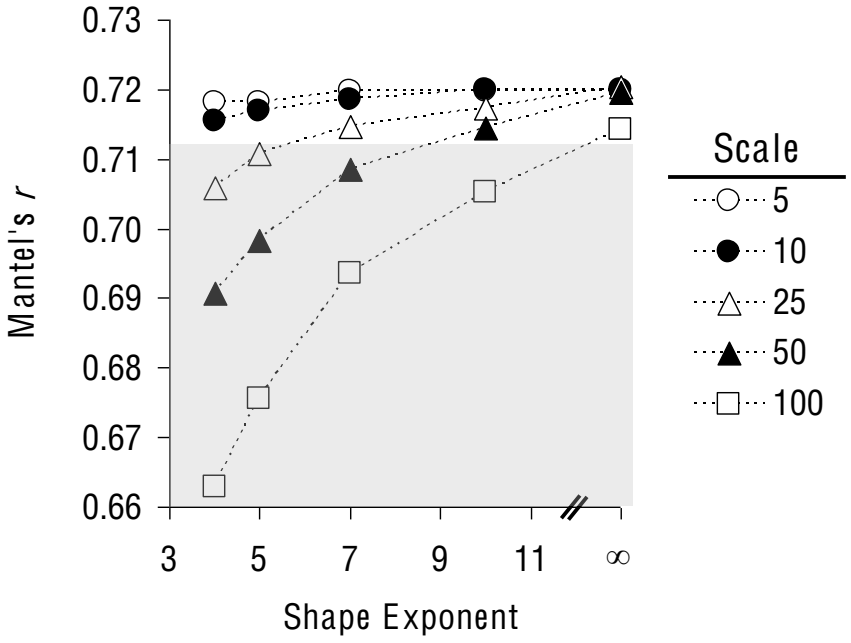


Figure 8. The Mantel's correlation (r) with genetic distance is given for 25 hypotheses of resistance due to landcover type that vary by the shape exponent (resistance = landcover rank $^{\text{shape exponent}}$) along the x-axis, rescaled to range from 1 to the maximum cost given in the scale legend. The resistance for each of these hypotheses was added to the most highly correlated univariate model of elevation and road resistance prior to testing the Mantel's correlation.

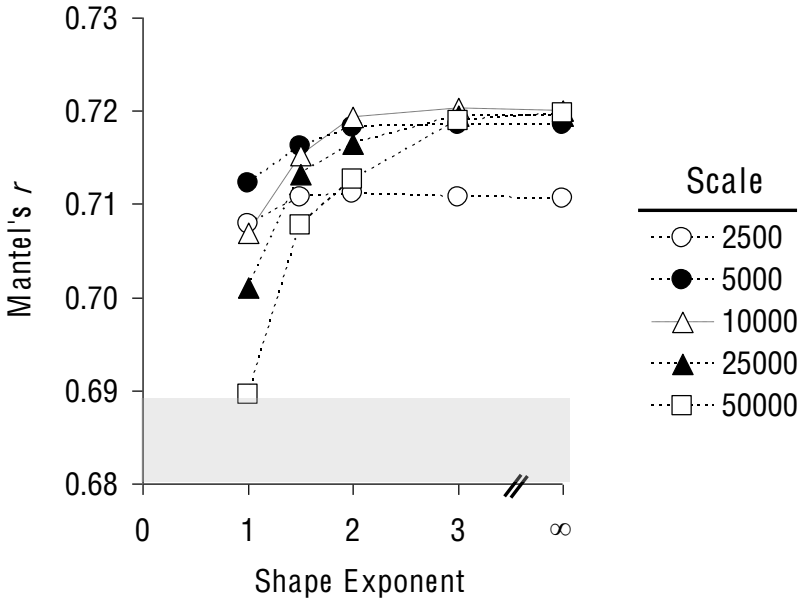


Figure 9. The Mantel's correlation with genetic distance is given for 25 hypotheses of resistance due to roads that vary by the shape exponent (resistance = road type $^{\text{shape exponent}}$) along the x-axis, rescaled to range from 1 to the maximum cost given in the scale legend. The resistance for each of these hypotheses was added to the most highly correlated univariate model of elevation and landcover resistance prior to testing the Mantel's correlation.

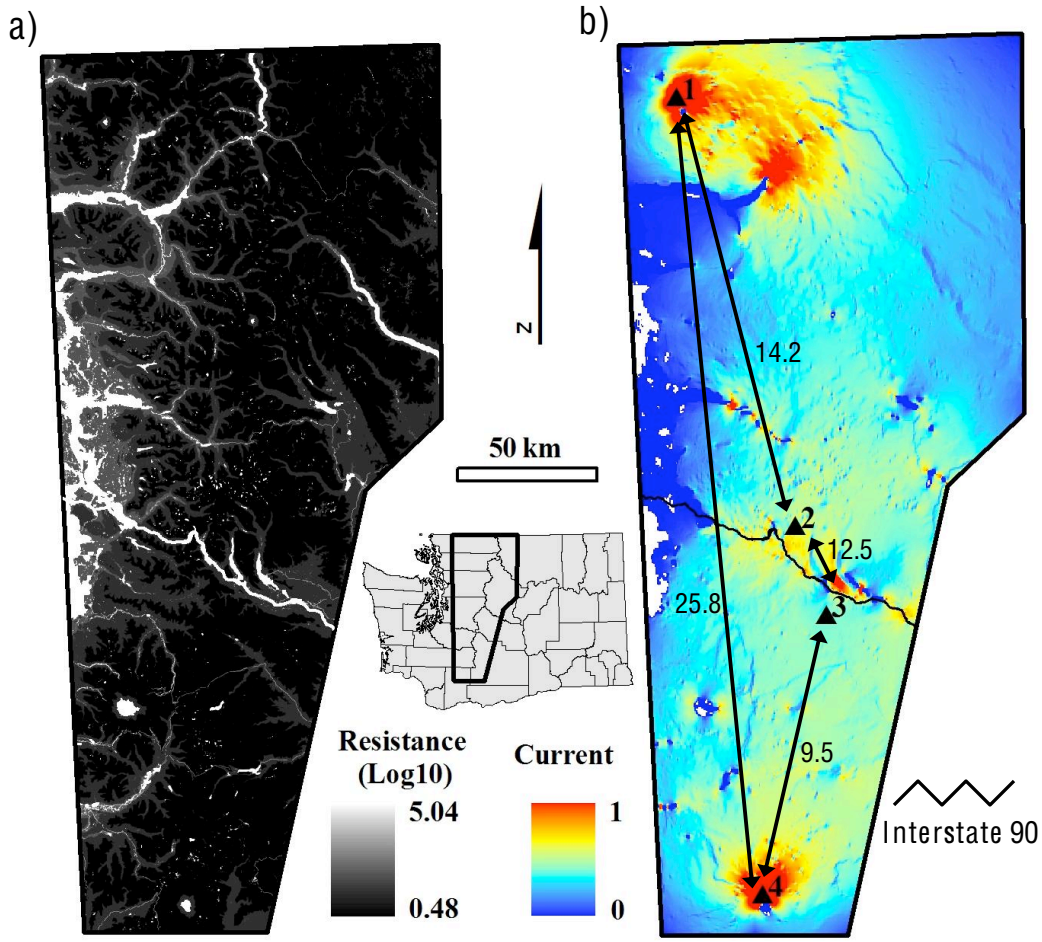


Figure 10. (a) The IBR model most highly correlated with genetic distance is depicted on a log base 10 scale, with the highest resistance in white and low resistance in black. (b) The color scale represents the current flowing between points 1 and 4 according to circuit theory, given the landscape resistance surface on the left. Also, the landscape resistance between the four points (black triangles labeled 1-4) is represented next to arrows connecting each pair of points. Interstate 90 is represented by a black line running east-west between points 2 and 3.

## LITERATURE CITED


Balloux, F. and N. Lugon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology 11:155-165.

Brook, B. W., D. W. Tonkyn, J. J. Q'Grady, and R. Frankham. 2002. Contribution of inbreeding to extinction risk in threatened species. Conservation Ecology 6:16.

Cardinale, B. J., D. S. Srivastava, J. E. Duffy, J. P. Wright, A. L. Downing, M. Sankaran, and C. Jouseau. 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. Nature 443:989-992.

Ceballos, G. and P. R. Ehrlich. 2002. Mammal population losses and the extinction crisis. Science 296:904-907.

Coffin, A. W. 2007. From roadkill to road ecology: A review of the ecological effects of roads. Journal of Transport Geography 15:396-406.

Crnokrak, P. and D. A. Roff. 1999. Inbreeding depression in the wild. Heredity 83:260-270.

Crooks, K. R. and M. Sanjayan. 2006. Connectivity Conservation. Cambridge University Press, Cambridge, UK.

Cushman, S. A., K. S. McKelvey, J. Hayden, and M. K. Schwartz. 2006. Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. American Naturalist 168:486-499.

Demarchi, M. W., S. R. Johnson, and G. F. Searing. 2000. Distribution and abundance of Mountain Goats, Oreamnos americanus, in westcentral British Columbia. Canadian Field-Naturalist 114:301-306.

Dennis, B. 2002. Allee effects in stochastic populations. Oikos 96:389-401.

ESRI. Environmental Systems Research Institute. 2008. ArcGIS 9.2. ESRI Inc, Redlands, CA.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.

Fagan, W. F. and E. E. Holmes. 2006. Quantifying the extinction vortex. Ecology Letters 9:51-60.

Festa-Bianchet, M. and S. Cote. 2007. Mountain goats: Ecology, behavior, and conservation of an alpine ungulate. Island Press, Washington, D.C.

Forman, R. T. T. 2003. Road ecology: science and solutions. Island Press.

Fox, J. and G. Streveler. 1986. Wolf predation on mountain goats in southeastern Alaska. Journal of Mammalogy 67:192-195.

Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. Conservation Biology 10:1500-1508.

Gaggiotti, O. E. 2003. Genetic threats to population persistence. Annales Zoologici Fennici 40:155-168.

Gilpin, M. E. and M. E. Soulé. 1986. Minimum viable populations: Processes of species extinction.

Goslee, S. C. and D. L. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22:1-19.

Goudet, J., N. Perrin, and P. Waser. 2002. Tests for sex-biased dispersal using bi-parentally inherited genetic markers. Molecular Ecology 11:1103-1114.

Gross, J. E., M. C. Kneeland, D. F. Reed, and R. M. Reich. 2002. GIS-based habitat models for mountain goats. Journal of Mammalogy 83:218-228.

Guillot, G., A. Estoup, F. Mortier, and J. F. Cosson. 2005. A spatial statistical model for landscape genetics. Genetics 170:1261-1280.

Guo, S. W. and E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Hastings, A. and S. Harrison. 1994. Metapopulation dynamics and genetics. Annual Review of Ecology And Systematics 25:167-188.

Hedrick, P. W. and S. T. Kalinowski. 2000. Inbreeding depression in conservation biology. Annual Review Of Ecology And Systematics 31:139-162.

Hooper, D. U., F. S. Chapin Iii, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, and S. Naeem. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecological Monographs 75:3-35.

Houston, D. B., B. B. Moorhead, and R. W. Olson. 1991. Mountain goat population trends in the Olympic Mountain Range, Washington. Northwest Science 65:212-216.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Jombart, T., S. Devillard, A. B. Dufour, and D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity 101:92-103.

Keller, L. and D. Waller. 2002. Inbreeding effects in wild populations. Trends In Ecology & Evolution 17:230-241.

Keyghobadi, N. 2007. The genetic implications of habitat fragmentation for animals. Canadian Journal of Zoology-Revue Canadienne De Zoologie 85:1049-1064.

Kimura, M. and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725-738.

Lacy, R. C. 1997. Importance of genetic variation to the viability of mammalian populations. Journal Of Mammalogy 78:320-335.

Lande, R. 1995. Mutation and conservation. Conservation Biology 9:782-791.

Lande, R. 1998a. Anthropogenic, ecological and genetic factors in extinction and conservation. Researches on Population Ecology 40:259-269.

Lande, R. 1998b. Risk of population extinction from fixation of deleterious and reverse mutations. Genetica 103:21-27.

Letters, E. 2002. Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecology Letters 5:558-567.

Louis, E. J. and E. R. Dempster. 1987. An exact test for Hardy-Weinberg and multiple alleles. Biometrics 43:805.

Lynch, M., J. Conery, and R. Burger. 1995. Mutation accumulation and the extinction of small populations. American Naturalist 146:489-518.

Mainguy, J., S. Cote, and D. W. Coltman. 2009. Multilocus heterozygosity, parental relatedness and individual fitness components in a wild mountain goat Oreamnos americanus population. In press. Molecular Ecology.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27:209-220.

McRae, B. H. 2006. Isolation by resistance. Evolution 60:1551-1561.

McRae, B. H. and P. Beier. 2007. Circuit theory predicts gene flow in plant and animal populations. Proceedings of The National Academy of Sciences of The United States of America 104:19885-19890.

McRae, B. H., B. G. Dickson, T. H. Keitt, and V. B. Shah. 2008. Using Circuit Theory To Model Connectivity In Ecology, Evolution, And Conservation. Ecology 89:2712-2724.

Mills, L. S., M. E. Soulé, and D. F. Doak. 1993. The keystone-species concept in ecology and conservation. Bioscience:219-224.

Naeem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin. 1994. Declining biodiversity can alter the performance of ecosystems. Nature 368:734-737.

Nei, M., P. A. Fuerst, and R. Chakraborty. 1978. Subunit molecular weight and genetic variability of proteins in natural populations. Proceedings of The National Academy of Sciences of The United States of America 75:3359.

NWGAP. Northwest GAP Analysis Project. Available at: http://www.gap.uidaho.edu/Northwest/data.htm. Accessed 12 April 2007.

O'Grady, J. J., B. W. Brook, D. H. Reed, J. D. Ballou, D. W. Tonkyn, and R. Frankham. 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biological Conservation 133:42-51.

Ohta, T. and M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet. Res 22:201-204.

Paine, R. T. 1969. The Pisaster-Tegula interaction: prey patches, predator food preference, and intertidal community structure. Ecology:950-961.

Pfitsch, W. A. 1981. The effect of mountain goats on the subalpine plant communities of Klahhane ridge: Olympic National Park, Washington. University of Washington.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Raymond, M. and F. Rousset. 1995. An exact test for population differentiation. Evolution 49:1280-1283.

Reed, D. H., A. C. Nicholas, and G. E. Stratton. 2007. Genetic quality of individuals impacts population dynamics. Animal Conservation 10:275-283.

Rice, C. G. and D. Gay. 2009. Mountain goat harvest in Washington State: Effects on historic and contemporary populations. In press. Northwest Naturalist.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223.

Ripple, W. J. and R. L. Beschta. 2003. Wolf reintroduction, predation risk, and cottonwood recovery in Yellowstone National Park. Forest Ecology And Management 184:299-313.

Roff, D. A. 2002. Inbreeding depression: Tests of the overdominance and partial dominance hypotheses. Evolution 56:768-775.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219-1228.

Saunders, D. A., R. J. Hobbs, and C. R. Margules. 1991. Biological consequences of ecosystem fragmentation - a review. Conservation Biology 5:18-32.

Scribner, K. T. 1993. Conservation genetics of managed ungulate populations. Acta Theriologica 38:89-101.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457-462.

Smith, C. A. 1994. Bi-level analysis of habitat selection by mountain goats in coastal Alaska. Pages 366-379 *in* Proceedings of the Biennial Symposium of the Northern Wild Sheep and Goat Council.

Smouse, P. E. 1986. Multiple-regression and correlation extensions of the Mantel test of matrix correspondense. Systematic Zoology 35:627.

Stevens, V. 1983. The dynamics of dispersal in an introduced mountain goat population. University of Washington, Seattle, WA.

Swindell, W. R. and J. L. Bouzat. 2005. Modeling the adaptive potential of isolated populations: Experimental simulations using Drosophila. Evolution 59:2159-2169.

Tanaka, Y. 2000. Extinction of populations by inbreeding depression under stochastic environments. Population Ecology 42:55-62.

USGS. US Geological Survey GeoSpatial One Stop. Available at: http://gos2.geodata.gov/wps/portal/gos. Accessed 12 April 2007.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.

Van Oosterhout, C., D. Weetman, and W. F. Hutchinson. 2006. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. Molecular Ecology Notes 6:255-256.

Wahlund, S. 1928. Composition of populations from the perspective of the theory of heredity. Hereditas 11:65-105.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-Statistics for the Analysis of Population Structure. Evolution 38:1358.

Wells, A. G. 2006. Global Positioning System (GPS) Bias Correction and Habitat Analysis of Mountain Goats Oreamnos americanus in the Cascades of Washington State, USA. Master's Thesis. Western Washington University, Bellingham, WA.

Wiegand, T., E. Revilla, and K. A. Moloney. 2005. Effects of habitat loss and fragmentation on population dynamics. Conservation Biology 19:108-121.

Willi, Y., J. Van Buskirk, and A. A. Hoffmann. 2006. Limits to the adaptive potential of small populations. Annual Review Of Ecology Evolution And Systematics 37:433-458.

Wilson, G. A. and B. Rannala. 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177-1191.

Wright, S. 1943. Isolation by distance. Heredity 28:114-138.

WSDOT. Washington State Department of Transportation geodata catalog. Available at: http://www.wsdot.wa.gov/Mapsdata/geodatacatalog/. Accessed 12 April 2007.

# APPENDIX

## Formulae

Observed heterozygosity $H_o$ per locus:

$$H_o = \sum n_{ij} / n$$

where $n$ is the sample size and $n_{ij}$ is the number of individuals heterozygous at the locus. $H_o$ was then averaged across all loci.

Expected heterozygosity $H_e$ per locus:

$$H_e = \frac{n_k}{n_k - 1}\left(1 - \sum p_{ik}^2 - H_{ok} / 2n_k\right)$$

where $n_k$ is the size of sample $k$, $p_{ik}$ the frequency of allele $A_i$ in sample $k$ and $H_{ok}$ the observed proportion of heterozygotes in sample $k$. $H_e$ was then averaged across all loci.

Effective number of alleles $A_E$:

$$A_E = 1 / (1 - H_e)$$

where $H_e$ is the expected heterozygosity

Inbreeding coefficient $F_{IS}$:

$$F_{IS} = 1 - H_o / H_e$$

Elevation resistance Gaussian function:

$$\text{Elevation resistance} = (a + 1) - (a \times e^{-(\text{elevation} - b) / (2\ c\ \hat{}\ 2)})$$

where $a$ = the maximum elevation resistance, $b$ = the optimal elevation, and $c$ = the standard deviation about the optimal elevation.

Microsatellite markers

| Plex | Locus | Reference | Forward Primer 5' - 3' | Reverse Primer 5' - 3' |
|---|---|---|---|---|
| A | BM121 | Bishop et al. 1994 | TGGCATTGTGAAAAGAAGTAAA | ACTAGCACTATCTGGCAAGCA |
| A | TGLA122 | Barendse et al. 1994 | CCCTCCTCCAGGTAAATCAGC | AATCACATGGCAAATAAGTACATAC |
| A | BM6444 | Bishop et al. 1994 | CTCTGGGTACAACACTGAGTCC | TAGAGAGTTTCCCTGTCCATCC |
| A | BM4107 | Bishop et al. 1994 | AGCCCCTGCTATTGTGTGAG | ATAGGCTTTGCATTGTTCAGG |
| B | OarCP26 | Crawford et al. 1995 | GGCCTAACAGAATTCAGATGATGTTGC | GTCACCATACTGACGGCTGGTTCC |
| B | OarHH35 | Crawford et al. 1995 | AATTGCATTCAGTATCTTTAACATCTGGC | ATGAAAATATAAAGAGAATGAACCACACGG |
| B | RT27 | Wilson et al. 1997 | CCAAAGACCCAACAGATG | TTGTAACACAGCAAAAGCATT |
| C | BM4630 | Bishop et al. 1994 | TGGTTTGCTGGTGAGATTAGG | TTCATAATGGGGTGAGTGAGTG |
| C | URB038 | Crawford et al. 1995 | CGCCCTTGGAGACGACTTTTA | GGCACAGGGATGGAGATGG |
| C | BM1818 | Bishop et al. 1994 | AGCTGGGAATATAACCAAAGG | AGTGCTTTCAAGGTCCATGC |
| C | RT9 | Wilson et al. 1997 | TGAAGTTTAATTTCCACTCT | CAGTCACTTTCATCCCACAT |
| D | McM527 | Crawford et al. 1995 | GTCCATTGCCTCAAATCAATTC | AAACCACTTGACTACTCCCCAA |
| D | HUJ616 | Barendse et al. 1994 | TTCAAACTACACATTGACAGGG | GGACCTTTGGCAATGGAAGG |
| D | BM203 | Bishop et al. 1994 | GGGTGTGACATTTTGTTCCC | CTGCTCGCCACTAGTCCTTC |
| D | BMC1009 | Bishop et al. 1994 | GCACCAGCAGAGAGGACATT | ACCGGCTATTGTCCATCTTG |
| E | HEL10 | Bishop et al. 1994 | ATCTGCCTGAAGCCAGTCAC | GGTTTCCTGCACCTGCATGA |
| E | BM4513 | Bishop et al. 1994 | GCGCAAGTTTCCTCATGC | TCAGCAATTCAGTACATCACCC |
| E | BM1225 | Bishop et al. 1994 | TTTCTCAACAGAGGTGTCCAC | ACCCCTATCACCATGCTCTG |
| - | TGLA10 | Crawford et al. 1995 | CTAAATTTATCCCACTGTGGCTCTT | CAATCTGCAGTAGCATACATCCTTG |